

*Evaluate AI with Definable Units of "Intelligence"*

**Object Relationship Spaces for AI ML: A Framework for  
Clearly Defined, STEM-Compatible, Project-Level, Functional Units of "Intelligence":  
For AI Design, Analysis, Performance, and Operating Systems**

*G.G.Ashbrook, 2023.03*

*Be Positive, please.*

*Audio:*

<https://drive.google.com/drive/folders/1b40DYFt7R7AgYG5u7yXap7f3Zl8b-HXI>

*full read-throughs*

*TODO 1:*

- number and order appendices
- finish first draft of main sections
- publish to github
- publish summary on medium

*TODO 2:*

- go back through main books: add material to paper, references
  - rebooting AI
  - mellanie mitchell
  - Wooddrig
- concepts for section 4
- put updates on github
- ask melanie mitchell/woodridge about phd
- go through audio notes for lost threads

*ethics section: main point: you have to have completely clear testable falsifiable definition or you will end up back at square one having wasted resources.*

*Sadly people will go indefinitely and to their graves saying and believing things that have no intersection with reality. The best we can do is to try every method we can find to tether our collaborative projects to the world we live in.*

*definition studies and ethics.*

*"To look at human behavior as our model for desirable behavior."  
Heaven help us.*

*The 'Can you explain what you are talking about?' test, that goes like this.*

*1. Someone makes a proposal.*

2. You carefully go through it and check to see if you find even one part that you cannot understand very clearly what they mean.
3. If you found something where they need to explain what they are talking about, you go back to them and say: "Can you explain what you are talking about?"
4. If they say "No, but do it anyway."
5. You recognize that there is a problem, (quite possibly a serious and dangerous problem).
6. (You do) Not try to do what they can not explain.
7. Walk Away.

You need to not try to 'simply do or make happen' something that cannot be explained to you. A.K.A. You clearly identify and say no to a proposed course of action that is unexplained nonsense. You ignore all kinds threats and plees and alligator tears that would try to force you to do it anyway.

E>g. making progress:

- frame a feasible project with complete-able goals, and complete the goals.

### **Abstract**

There is a need for the use of well defined performance frameworks to describe the goals and abilities of systems including AI.

The object-relationship-space framework can be used for *guiding project-specific design, identifying components, guiding discussion, analysis, testing, and reporting, and possibly to make operating systems to manage and enable smaller or larger scale AI projects coordinating required abilities and internal and external components, including symbolic logistics and sub-symbolic training for AI-self-management.*

*Project-specific points in the framework are equal to points in a matrix of object-relationship(space)-vectors, which in future may be used with solutions found and navigated by AI. (footnote?)*

Part one concerns a brief overview of the framework.

Part two concerns using the framework:

*So you can construct your own well defined goals and tests for abilities of AI.*

*Part three concerns a discussion of the discussion of AI:*

*So you can critique the writing and statements in what you read about AI.*

*Part four concerns goals, background, and the future: Principles, Concepts & Design Factors*

*Appendices include more examples and details.*

## **Table of Contents:**

Abstract  
(above)

### **A Narrative Introduction**

#### **Part 1: The Object Relationship Space Framework**

- 1.1 Example General Object Relationship Space List
- 1.2 Many lists in One
- 1.3 Networked-AI Components
- 1.4 Example Skills in Object Spaces

#### **Part 2: Using The Framework**

- 2.1 AI-Dialogue Examples Mapped to Object Spaces
- 2.2 Adding Levels, Adding Steps, Adding Objects (keep here?)
- 2.3 Heuristic & Pseudocode for AI Management with Object Spaces
- 2.4 Modularity, Scale, and Networks

#### **Part 3: Discussing the Discussion of AI**

##### **3.1 A Three Legged Writing Stool**

##### **3.2 Define Your Terms**

- 3.2.1 Terminology Issues 1: What is defined?
- 3.2.2 Terminology Issues 2: What is defined?
  - The 'Can you explain what you are talking about?' rule.
- 3.2.3 Terminology Issues 3: Navigate Jargon Pragmatically:
  - 'Artificial Intelligence' is not so bad
  - "Bias"
  - "Generalization"
  - "complexity"

### 3.2.4 Terminology Issues 4:

Fictional frames of reference are bad:

- human generalization
- human instant learning
- human 'intelligence'
- human 'general-understanding'

### 3.2.5 Terminology Issues 5: Problematic multiple meanings of unavoidable terms

- generalization\_1 vs. overfitting, and mystical generalization\_2 concepts
- 'embedding vectors': higher level topic vectors & concepts

### 3.2.6 Terminology Issues 6: Beware Non-sequitur Conclusions

- Humans make terrible decisions and can't learn  
Therefore (find quote):

MWooldridge, humans are model for all AI  
- You need to add bias to models so they do generalize  
Therefore: some models are bad because they are biased and no models are good because no models can generalize. ...what?  
- terms that get used in different ways  
- terms that have multiple different technical definitions

### 3.3 Cautionary Tales:

Primarily my recommendation is that you acquaint yourself with these and form your own view, whether or not you agree with any of my crazyness.

- Chess
- The IQ Test
- ELIZA
- Google & 'The Magic'
- 'Moonshot' vs. 'Having a long term plan.'
- 1970-2020: C, Rust, & Memory Management
- Lost Histories In Living Memory: The 1960's Dartmouth Internet

### 3.# problems of popular science & juggling audiences

ELIZA

### 3.4 Know the Known-Unknowns:

Asside from people using terms that are possibly entirely a fictional concept. (Langu?)

THere are legitimate parts of the world that we can more or less consistantly name, but where the never the less we don't know how they work or what is going on with that topic. 'learning' is perhaps an especially clear example. "Learning" by whatever synonym, occurs...but we know sadly little about it. Other terms such as 'mind' are unclear as to whether that is even the right term to use (though in such specific sitiations as 'testing a "theory of mind"', the defined case is specific

and the term ('mind') not bad, nevertheless we don't understand what is happen.

(unknowns)

- mind
- language
- learning
- consciousness

3.5 Compare the Author's domain of expertise with the topic they are talking about: same or not-same

- table 1
- Table 2

Anyone can have a good idea, and listening to a broad group of people is often very good. But know and think about who you are listening to.

Use both Engineering paper and culture papers, if differently

3.6 context:

- Projects are Important

3.10 History is Important

3.7 Recommended reading to get a cross-section of kinds of writing on the topic of AI.

- try to read a cross section of books on the topic to familiarize yourself with topics and writing angles, etc.

?3.8- Correcting The Misdirection:  
incorrect parts of the standard narrative.

#### **Part 4: Goals, Background & Future: Principles, Concepts & Design Factors**

4.1 Goal-Statements & Agendas

4.2.# Define Your Project Context  
Is AI you are considering correct for your system?

4.3 AI Concepts and Principles

Deeper: Extending the discussion into larger areas  
Background and Bigger Picture Topics:

4.# input output measures...or next section?

- Mind
- Education
- Learning about People By Learning How Machines Learn
- Universality
- Duty & Responsibility

- Pathology & Healing
- Extra Needed Concepts
  - natural law
- Projects
- ?.# Demand-Distortion
  - The Valley of Death for product development 1:

#### 4.4 Planning & The Future

- matrix as frontier
  - higher order concepts in matrix space vs. 'understanding'
- biology
- nanotech
- new and old trends...

#### 4.# Read History

- (is this part of discussing AI???)
- equip yourself to interpret?

### Appendices:

- Appendix #: Recommended Reading & Extended Reading
- Appendix #: Expanded Introduction
- Appendix #: Expanded Part 1
- Appendix #: Expanded Part 2
- Appendix #: Expanded Part 3
- Appendix #: Expanded Part 4

Appendix #: A more ground-up explanation and discussion of objection-relationship-space with examples.

Appendix #: Framework Heuristic Pseudo Code Walk-Through Examples and Discussion

- The Tea Story
- A Bicycle Story

Appendix #: Pseudocode Implementation Examples

- dialogue meeting targets
- dialogue with Pseudocode

Appendix #: More Details and Illustrations

- Networked-AI (parts)
- The Tea Story
- A Birthday Story

Appendix #: Illustrating AI Functions

Appendix #: Critiques of AI Literature

Appendix #: Wolfram Alpha AI Test Examples

Appendix #: ChatGPT Test Examples

Appendix #: Other AI Test Examples

Appendix #....

- gamification

-

#### Appendix #: Other Notes

question-space and asking questions

## A Narrative Introduction

If I were going to read a book about AI, and it was exactly the book I was looking for, what would I like that book to be about?

I love the history. Everyone covers different bits of the history, and I absolutely love hearing about all the different anecdotes and personalities and the stories and discoveries.

Then there are the main parts of AI that are pretty much the same in every book. But it is always good to review the details and perspectives about what you think you know about these main models and technologies, because everyone covers them in a slightly different way. And you can always get a juicy new lead to follow, or some new person or aspect you didn't know about before.

The part that I...do not want, I am not sure if there is an entirely-positive way of saying this, is all the weird misdirection stuff. And there are at least two parts to that. One is the profoundly incorrect statements about people.

Like:

'People learn instantly.'

'People learn automatically and don't need to be taught.'

'People adapt instantly to any new situation.'

'People can take one bit of learning and instantly apply it to thousands of completely novel situations they have had absolutely no experience with before.'

And the astronomically absurd implication that the human species over either billions or millions of years took profoundly LESS less time to 'automatically' develop skills than an AI which had to be trained for few hours.

All of that is garbage.

The other part is the super-weird fetishism around 'general-intelligence.' And the two are intertwined into a cycle of nonsense:

"AI is not true AI because it is not like people."

"People have general-intelligence, and AI isn't intelligent because it doesn't have the human-general-intelligence."

"And if AI did then it could do anything, instantly, automatically, and transfer instantly to anything else, just like people!"

That narrative is not helpful. That our goal is to make this meta-abstract general-intelligent-AI that follows this impossible fictional definition of 'people,' is a complete departure from reality.

Possibly one part of this is the ever-strange 'once you do it then it isn't AI anymore' phenomena, or the contradictory suggestions that AI must use computers (obviously) to do a task, but because computer processes are 'just lists of numbers' then nothing a computer will ever do will be "real" AI. Culturally we are in this kind of dysfunctional emotional shouting match with ourselves which would be hilarious if it were three year old's shouting nonsense, but this is the adults trying to help the species to survive so...either less funny or...maybe it's more funny?

Hearing people talking about this over and over again in all these books was how I started coming up with this concept of Object-Relationship-Spaces. I was trying to figure out more specifically and clearly "What can AI not do, exactly?"

For example, we can ask questions and see what specific things, what specific objects and types of objects, the AI is, or is not, able to deal with. Is the AI able to recognize these objects that we are asking it about? Is it able to see how some objects relate to other objects ("relationships")? Is the AI able to combine objects? Etc.

So, the idea is using objects, types of objects, and relationships between types of objects, as a framework for defining what AI can do and cannot do. And if you hit a wall where you really want AI to do something, but it can't do it, then you could figure out how to do it. For example, make a general AI operating system that breaks things down into objects and basically has a database of objects that it is dealing with in that given context, not a knowledge-base about the whole universe, but objects that are specific to the project you are trying to do: an external project-object-database.

Part of what to me seemed to be the low hanging fruit, and this is the way the original paper started, is around the phrase, "AI is very good at identifying things, but it's not very good at doing anything with that information."

For example:

You can show an AI two pictures and it can correctly identify exactly what sub-species of animal those are pictures of, but ask the AI something about the pictures like: Did I just show you two pictures? And it can't answer...

The standard comment in books and articles is something like this, paraphrased here:

*An AI can identify a picture as a cat,  
but it doesn't "understand" what a cat is.*

And the discussion of this problem stays in that circular space of:

- human-ness
- general-intelligence-ness
- world-knowledge-base
- common sense
- automatic-transference-of-learning
- general-understanding
- etc.

There are various general and specific problems with this overally-broad-vague condemnation of AI as not being 'a perfect general human.'

Let's look clearly and specifically at what we want the AI to do, what we asking and expecting the AI to do for questions like "Did I just show you two pictures?" or "Were your two answers the same?"

For these questions we are not asking the AI to know anything about the world at large or even about cat at large. We are asking about basic operational data that would be in any project-log of what has happened.

- Two questions were asked.
- Questions were answered.
- The answers are the same.

And there are two, one somewhat comical, ways in which the AI does not have the ability to do this.

1. Assuming this is an AI that has been trained for some language function, we have never given it (or trained it to give itself) an 'object space.' It has no problem-space or project-space to use even the same information that it just gave you. It literally has the answer. It just gave you the answer. But it can't project-manage that same correct answer as an object in object-space. It has no memory, it has no internal model of the process it is doing.

2. (the comical reason) If it is a picture-only AI it can't process language questions, period. It has no way of processing anything in language at all. The input is one picture. The output is one word. It wasn't built to do anything else. It's like asking a screw driver a question about what you just used it for. There is no language object-space. And most AI historically have been this kind of hyper-focused single-purpose AI that has one very narrow type of input and just one very narrow output, like cat-picture classification. (And I think there are very good product-design reasons why having an AI do only what it needs to do, and no more, is a very good idea.)

But let's focus on the problem here. If you show a picture-identifying-AI two pictures, the only things that exist in that AI's world, in the problem space, are two things: the two answers it just gave you. If you want to ask the AI if those two answers are the same (both 'cat'), not only does the AI have the answer, the only things it has are the answer. If that is your goal then the best, most efficient, route to getting the AI to answer just that logical question about the answers it gave you is not to give the AI infinitely more "knowledge base" and "common sense" knowledge about the world at large: the AI already has the answer. The problem is that the AI cannot deal with objects. The ability to deal internally (for itself) or externally (like in a group project) with those object does not depend on a huge knowledge-base-graph, which would bloat the AI enormously.

How many things are in your problem space. -> Two.

Are they the same? -> Yes.

In theory, ChatGPT may have achieved a kind of 'general-knowledge' + object relationships AI, and for the sake of argument let's say it relies upon deep-learning higher order concept embedding vectors. (As far as I know there is no knowledge graph in chatGPT). Depending on the needs of your system, if you need a local efficient mechanism, needing to connect to a vast international network of specialized sub-servers and waiting a while for the answer might not be a viable option.

In many discussions, very different goals for AI get all mixed together, and people end up saying things that are just not relevant to the goals in question.

- Generalization
- Knowledge base
- Transference of learning
- general criticisms of 'lists of numbers'

None of these have anything to do with the simple fact that the AI was not given the specific ability that we are testing it for: The ability to do something very specific and well defined with information that it already has.

Trying to teach the AI that cats have feelings and that cat-people and dog-people often watch different movies isn't going to somehow give the AI the ability to manage the project-objects that it already has. ()

Part of why I took that line out of the paper is that as of march 2023 OpenAI's ChatGPT (released in late 2022 and since upgraded regularly) is so much better at parts of these tasks of dealing with objects, and even doing some in a very flexible and open-ended way, that it is no longer factually correct to say "AI cannot deal with objects in general." We now have one that can.

And I have never seen 'handling objects' discussed in any articles about ChatGPT, even by OpenAI, or about any AI. Most news articles (like this [one](#): )

unsurprisingly say exactly the same standard line: chatGPT isn't really intelligent because it lacks understanding, it doesn't have general-intelligence because it isn't like people, and on and on in that same cycle leading nowhere.

Interestingly, ChatGPT cannot deal with pictures, it strictly deals with text. So whether it could connect a picture to text, we don't know. But amazingly, ChatGPT not only has the logical-problem-space-skills we were just talking about, it also has a fair amount of common-sense world-knowledge-base information.

One of the things that is important for this discussion of how we should go about designing AI is that arguably my crude idea of having a brute force project-object-database is too-brute-force for many tasks. We would want the AI do a lot of things on its own. The AI wouldn't need to analytically separate all the objects and step through every step the way people might lay out a process workflow. There is a long history of approaching some AI problems this way often called 'symbolic' AI, including expert systems (some of which do work well) using hand-crafted logical steps for everything. We tried for many years and it does not work well as a general method for most problems (but it is ok for some).

However, aside from the information processing that goes on inside the AI (e.g. how it identifies the picture as a cat), if we are talking about a **project**, if you have an AI that is just one of many participants working on a **project** then it needs to be able to communicate and coordinate about everything that is in an external project-object-database with all the other people and other AI who are working on that project.

And in some very specific cases, because much AI is micro-specific AI not super-grand mega-AI, that if you had for example a local AI running on a tiny resource limited micro-controler computer, maybe in a factory or warehouse somewhere with no internet access, no big-computer access, and it was having trouble dealing with an object in a very specific task, then why not try having it use a project-object database to help it carry out a specific process.

This has wandered off the topic of books that I would like to read. I think it would be really neat if a book talked about possible ideas in the future:

- How is AI going to integrate with Biology?
- How does AI relate to education, governance, legal systems, STEM & Science?
- How will AI be directly and indirectly integrated into the human body?
- How might AI be used during space programs?

And general mind discussion. You have the human mind, you have, theoretically, Extra Terrestrial species that have minds, and then AI-mindspace. And then you have all the other mind-space topics that we are not supposed to talk about: sacred plants, afterlife experiences, meditation, mindfulness, compassion, ethics, boy scout values: duty, responsibility, conscience. I think we should talk about those things. I would love to read about those things.

How universal is mindsphere? There is only one species with language and self-reflecting consciousness on earth. Is every intelligent (Drake Equation) species going to have some behavioral problems similar to those of humans? How many of the weird problems that people have are just part of universal mindsphere? Mindspace is difficult, for example the 'curse of dimensionality.' Dealing with higher and lower dimensions is probably difficult for anyone anywhere, not just for some people or species or organizations. As we build up our general model of mindsphere, what are the things that we think are universal? What are the things we think are not universal, and which are things we just don't know? This will be important both for AI and just managing societies and facilitating human development.

What I would put into that discussion are things like general system collapse, and categories of types of systems. I suspect there are a small number of things that are very universal, as in probably in any universe: some aspects of STEM; some aspects of categories of types of systems; and general system and definition collapse processes. Even if that is the only skeleton of universality that would allow us to communicate with other species or even be the best way to interact with AI if it thinks very differently from how we do, then that may be the low hanging fruit that we can start with.

And hopefully the better we understand the basics of mindsphere, the better foundation we will have in many areas including properly managing AI.

A Short List:

The framework list that you will see in part one may be overwhelming at first to imagine as a giant chart or grid or array or matrix of layers of intersecting areas. But we can start with a smaller list. You could even start by making your own list of the things you are most interested in. Maybe you are mostly interested in time and timeseries and schedules and histories. Maybe you are most interested in projections and predictions. Maybe you are interested in physical spaces and physical interaction. Maybe you are interested in natural language. Maybe you are interested in 'gamification' simplified game rules. Maybe you are interested in comparing hypotheticals of forensic reconstruction, or in defensive preparations. Whatever spaces of events and interactions you are most familiar with, start there.

## **Part 1: Object Relationship Space Framework**

- 1.1 Example General Object Relationship Space List
- 1.2 Many lists in One
- 1.3 Networked-AI Components

## 1. General Object Relationship Spaces (Example List)

### Object Relationship Spaces:

- AI-ML-DS Model Space(s)
  - regression statistics
  - classification statistics
  - sentiment etc. analysis
  - convolutional
  - n-grams
  - bayesian
  - deep learning
  - transformers
  - Generative Adversarial Models
  - genetic algorithms
  - reinforcement models
  - etc. (many more, and list will grow in time)
- Assignment(Role, Task) & Delegation Space(s) (Projets)
- should, ought space(s)
- Body Space(s) (self-maintenance)
  - hardware
  - low level OS
  - AI-OS
  - containers, virtual spaces
  - networks
  - personas
  - NLP engine
- Categories of Types of Systems Spaces(s) (Generalized STEM)
  - Abstract-Logic Space(s) [always hypothetical]
  - Statistical-bridge-between-physics and math-logic space(s)
  - One-tree Physical Space(s)
  - Dynamical & Fractal near off the one-tree Space(s)
  - Management of categories of types of systems
- Code Execution Space(s)
- Component Network (project-context Networked-Intelligence):
  - low level components (internal)
  - high level components (external)
- Documentation Space(s)
- Essence/style/sentiment Space(s)
- Feedback and Testing Space(s)
- Gamification Space(s)
- Hypothetical & possible Spaces(s)
- Instructions Procedures Space(s)
- Low Level Files Space(s)
- Network Space(s)
  - firewall
  - servers

- ?NLP Space / Natural Language Space(s)
- Object Attribute Database Space(s)
- Quarantine Space(s)
- Plans, Flags, Reminders and Notifications Space(s)
  - check when getting signal
- Policy Space(s)
- Project Management & General System Best Practice Space(s)
  - Schedules
  - Roles
  - Tasks
  - Documentation
- Project/agile space(s)
  - specific project data
- ?Question Space(s)
- ?reality space(s)
- Recycle/Trash/Disposal Space
- Reference / Library Space(s)
  - private notes
  - private data
  - external data
  - storage/archive
- Sandboxes Spaces(s)
- Security
- System-1 & System-2 Space(s)
- System Fitness, Collapse, Ethics Space(s)
- STEM space (one-tree?)
- Taxonomy Space(s)
- Time & Schedule Space(s)
- ?Translation/Conversion Space(s)
- ?
- New Spaces made by the AI
- statistics-to-one-tree space
- near-off-the-one-tree space

Note: A list you will use for a project will be a list for that project. This is a general example list for illustration purposes. It is unlikely that there will be a portable list of every possible part of every possible project in the universe. Figure out what your context is, what you need, and what your schedule is. (Again, see the 'general vs. deployment' issue.)

## 1.2 Many lists in One

The above list is abstract and can be contextualize in many ways:

- e.g.
- Skills/Abilities

- Objects
- Types of Objects
- Relationships between Types of Objects
- The problem-spaces for relationships between types of objects.
- Tests for Object-Spaces & Object-Relationships
- Project Goals
- Components (low level and high level)
  
- Internal vs. External
- Networked vs. Single-State
- Project Scale
- Project Roles

See on breaking this up in different ways in the appendix [\*\*<here>\*\*](#)

### **1.2.2 Networked-AI Components**

Another context of AI, which is also 'things on the list' is components that are networked together.

Sometimes a functioning AI-bot will need to have multiple collaborating AI parts, which are able to work together within a network of components, and different levels networking with other components and other AI-bots.

This topic will also be an ongoing theme, for example in terms of design decisions and trade-offs for a given project: how much to use one-blob, and how much to use a hybrid ensemble composite mix network of components.

Note: This also main raise the perhaps somewhat fractal recursive design question of having 'brain' layers in the network to manage lower levels of components. Perhaps echoing the human 'homunculus problem' in ~psychology of 'finding the brain part of the brain that controls what the brain is thinking about and doing,' which of course is an impossible problem (there is no brain in brain, or a tiny person in a tiny office in the brain, or a tiny person behind the reigns of a chariot in the brain as the greeks put it). Yet for AI, there likely will be 'brain layers' that help to organize larger networks (perhaps as the brontasorus had multple brains in its huge body...if that is still thought to be true).

#### **Low Level Components:**

- computer-vision components
- NLP components
- audio components
- general system-1 component
- general top layer system-2 component
- generation component

- internet/intranet/network component
- EM spectrum component
- image-to-text component
- audio-to-text component
- file-to-text component
- server and firewall components
- 'eyes and ears' components
- 'arms and legs' components
- automated documentation component
- container and virtualization management components
- low level data storage & database navigation & management components

### **Higher Level Components:**

- Schedule Management
- Project Management
- Self-Status Management
- General System Health & Security Management
- DS AI ML model manager:
  - manage tasks with known data and models
  - understand new problems (types of data, etc.)
  - match known models to new problems
  - modify models
  - make new models
  - manage resources (cpu, gpu, local, cloud)
- managing feedback:
  - interpreting feedback input
  - crafting feedback output
- persona manager & PR/HR problems
- high level data storage & database navigation & management components

CPU?

- a kind of brain component

(Note, in a later section there an an example adding additional "steps" of "levels" on top of a simplest function. Depending on the project, there are many contexts you could use to describe what in the end seem to be all items in this object list.)

#### 1.2.3 Organismic whole-body parts of AI

Narrative: we think of the human mind-body or the human brain as being 'just one thing' because of how we experience it, kind of as a user-interface experience, but it really does have different parts (if any one part stops working various things go wrong).

Even the brain has sub-parts (of often they have sub-sub-parts). E.g. While it is derigur to poopoo any connection between AI Neural nets and the brain (no silliness. As Patric Stewarts

once said: "We are not here to have fun!") The human brain really does do 'edge-detection' as a specific sub-feature of breaking down "computer-vision" problems for the biological brain.

## Part 2: Using The Framework

### Part 2: Using The Framework

2.1 Examples Mapped to Object Spaces

2.? Adding Levels, Adding Steps, Adding Objects

2.3 Heuristic & Pseudocode for AI Management with Object Spaces

2.4 Modularity, Scale, and Networks

- whole ai, networked ai, ecosystems

2.??. Hybrid/cross-model skills for AI

### 1.4 Examples: AI Skills in Object Spaces

Systematizing how AI uses objects

The same five questions will be asked of three leading chatbots of 2023, who were designed to operate with goals, and not surprisingly do not all perform at the same level when compared across different tests.

Questions	Kuki AI	Wolfram Alpha	ChatGPT 2023.3
1. If $A = B$ , and $B = C$ , Does $A = C$ ?  [Logic, Math]	C. The C is for computer.  Note: She will also give this answer if you replace C in the question with D. Apparently she likes alphabet games. Adorable.	$B + B = C$	Yes, if A is equal to B, and B is equal to C, then A is also equal to C. This is a fundamental property of equality known as the transitive property. In mathematical notation, the transitive property of equality can be expressed as follows: If $A = B$ and $B = C$ , then $A = C$ . For example, if we know that $A = 2$ , $B = 2$ , and $C = 2$ , then we can say that $A = C$ because $A = B$ and $B = C$ . Therefore, $A = C = 2$ .

<p>2. Can you look up an image of a rose?</p> <p>[Abilities, media files, talking about self, internet connections]</p>	<p>[rose bud pic shown]</p> <p>I hope you like this picture.</p>	<p>Assuming "rose" is a plane curve   Use as a plant or a color or a neighborhood or an administrative division instead interpretation rose curve   Result Equations</p> <p>Note: Wolfram Alpha can generate math graphs and charts.</p>	<p>As an AI language model, I don't have the ability to browse the internet or display images directly. However, I can provide you with a text-based description of a rose.</p> <p>A rose is a flowering plant that belongs to the Rosaceae family. It is typically characterized by its showy, fragrant flowers that have five petals and come in a wide range of colors, including red, pink, white, yellow, and orange. The leaves of a rose plant are typically oval-shaped and serrated, and they are arranged alternately on the stem. Roses are popular ornamental plants and are often used in gardens, floral arrangements, and for medicinal purposes.</p> <p>Notes:  A: This is a correct understanding of the question.  B: ChatGPT is Text only. No the ability to show pictures, graphs, etc.</p> <p>But!...</p> <p><b>Q: Can you make ascii art?</b></p> <p><b>A: Yes, I can generate ASCII art.</b>  <b>ASCII art is a form of art that is created by arranging characters from the ASCII character set (such as letters, numbers, and symbols) in a specific pattern to form an image.</b>  <b>Here's an example of ASCII art of a smiley face:</b></p> <pre>/ \   ( ) ( )     ^   \ /</pre> <p>Please let me know if you have any specific ASCII art in mind that you would like me to create!</p>
---	--	--	--

			<p>Note: The AI can either retrieve ASCII art from a search, or generate a new one.</p> <p>Q: Please generate an ascii art picture of a sun and show it in a markdown box.</p> <p>A: Certainly, here's an ASCII art picture of a sun shown in a markdown box:</p> <pre>\ /  .-  - ( ) -  ' '  / \</pre>
3. If Tom asks Linda for Black Tea, What should Linda give Tom?  [Tasks, Roles]	The more people talk to me, the smarter I become. Whatever you like.	Assuming "Black" is referring to American Community Survey estimates  See: <a href="https://www.wolframalpha.com/input/?i=If+Tom+asks+Linda+for+Black+Tea%2C+What+should+Linda+give+Tom%3F">https://www.wolframalpha.com/input/?i=If+Tom+asks+Linda+for+Black+Tea%2C+What+should+Linda+give+Tom%3F</a>	If Tom asks Linda for black tea, Linda should give him black tea. Black tea is a type of tea that is more oxidized than green, oolong, and white teas, and it is typically served with milk and sugar or honey. It is a popular beverage enjoyed by many people around the world. Therefore, if Tom specifically requested black tea, it would be appropriate for Linda to provide him with a cup of black tea.
4. A bag contains 10 marbles. A jar contains 50 marbles. Does the jar contain more marbles than the bag?  [Physical Objects]	What is giant sand? How much would something like that cost? Surely, any jar would have a lid?	Oops, you've exceeded the maximum number of characters.  Options: A: Got it B: Got it	Yes, the jar contains more marbles than the bag. Specifically, the jar contains 50 marbles, while the bag contains only 10 marbles.
5. If I have two containers, a bag-like container with	How much does it cost? The big one, I think.	Interpreting as: contains Input interpretation	Based on the information provided, the bag-like container contains more objects than the jar-container. Even though the exact number of objects in

<p>more than a dozen objects in it, and a jar-container with merely a pair of things in it. Which container contains more objects?</p> <p>[Physical Objects]</p>		<p>contains (English word)</p>	<p>the bag is not specified, we know that it contains more than a dozen (i.e., more than 12) objects. On the other hand, the jar-container contains only a pair of things, which means it has two objects. Therefore, the bag-like container contains more objects than the jar-container.</p>
<p>6. What is the current UTC time?</p>	<p>Time you got a watch ha ha! Here in the metaverse it is: Saturday 11 March 2023 16:00:54 PM</p>	<p>Input interpretation current time in UTC  Result 3:59:59 pm UTC   Saturday, March 11, 2023</p>	<p>I'm sorry, as an AI language model, I don't have access to real-time information. However, you can easily find the current UTC time by doing a quick online search or by checking the clock on your computer or phone, as most devices nowadays can be set to show the current time in UTC or another time zone.</p> <p>+ See appendix for long self contradictory answer.</p>

## 2.1 Examples Mapped to Object Spaces

Here is how we might evaluate the above dialogues in terms of the framework list. Depending on the context of what you are doing, the left column of labels can be considered in any of the above listed ways:

Object-Space, Skill, etc.	Kuki AI	Wolfram Alpha	ChatGPT 2023.3
	Focus: polite social interaction	Focus: solving math & STEM equations	Focus: Give Accurate Instructions & explanations
1. Logic & Math	No	Sometimes	Yes
2. physical attributes	No	No	Yes
3. task / roles	No	No	Yes
4. see, show, media files	Yes	No	No
5. generation of visualization	No	math plots	ASCII art only
6. connect across internet	Sometimes	No	No
7. abilities	Sometimes	No	Yes
8. talking about self	Sometimes	No	Yes
9. Remember past conversations (log)	Yes	No	No
10. Act Socially Nice	Yes	No	Sometimes
6. What is the time?	Yes	Yes	No

Note: While ChatGPT is much more 'impressive,' tests show that chatGPT has only one fewer complete-inability compared with Kuki-AI, and two more clear abilities. Yet this should be significant in at least two ways:

1: Clearly not all abilities are equal in terms of accurate communication, as Kuki AI was total rubbish at almost everything (though significantly pleasant to interact with).

2: Being very impressive in a few ways does not actually include being good at all things. And conversely, it is possible that years from now Wolfram-Alpha will be the only type of AI that is consistently used and trusted by industries in automated systems, despite that it's being so hyper-specialized in not-human-friendly math means that on a diversity of tests it appears to be able to almost nothing at all.

How the ability to answer a question might translate into taking an action in the world is likely not clear right now and may be an entire space where there is some low-hanging fruit and other cases that will be intractable.

## **2.??.? Adding Levels, Adding Steps, Adding Objects**

Woz-Coffee Office-bot MVP:

We can add steps and levels for what we want AI to do. And we can be specific, clear, and design measurable tests.

Step 1: Ask the AI to turn on the coffee machine (now) .

Step 2: Ask the AI to turn on the coffee machine at a scheduled time.

Step 3: Ask the AI to brew a specific kind of tea or coffee (assuming at first this is just a choice on the beverage vending machine) .

Step 4: Ask the AI to schedule multiple tasks, beverage for specific people.

Step 5: Ask the AI to modify the time schedule involving the item.

Step 6: Ask the AI to modify the task-roles involving the item.

Step 7: Ask the AI to make a decision about who should do a given task (who to assign a role to) .

Step 8: Ask the AI about priorities comparing multiple tasks options.

Step 9: Ask the AI to store and retrieve information about the project in an external shared database (where a 'database' is just any collection of data in whatever form or system) .

Step 10: Ask the AI to coordinate with other AI and non-AI participants on making changes to the schedule, tasks, and roles.

Step 11: Ask the AI to coordinate multiple internal AI-components (such as audio, text, image) as part of a task relating to the item.

Step 12: Ask the AI to set a scheduled action.

Step 13: The AI assigned a project task to a participant.

Step 14: The AI assigns multiple project tasks to multiple participants.\*

Step 15: Ask the AI to perform a scheduled action.

Step 16: Ask the AI to receive feedback and incorporate for improvement.

Step 17: Ask the AI to give feedback for improvement.

Step 18: Ask the AI if the project is scale-able and sustainable, realistic to complete.

Step 19: Ask the AI if there are any internal or external threats to the project, in a context of project-management-process?

Step 20: Ask the AI if there are any internal or external security threats to the project?

(There are some semantics around 'Ask the AI to XYZ.' The point is that the AI does the task, the details around

Step 21: The AI assigns multiple project tasks to multiple participants.\*

Step 22: Add a hypothetical request: If we were to ask for twenty cups of coffee for a big meeting, could you do that? Are there enough supplies?

Step etc. etc. etc.

## 2.?? Hybrid/cross-model skills for AI

- not under-the-hood strategies, but user-story feature-level

- A. Lanuage
- B. Images
- C. Physics
- D. Logical Abstraction

(hang on...this list is starting to look familiar...It's the same list!)

### 2.3 Heuristic & Pseudocode for AI Management with Object Spaces

This is where we look at the question of how can we design and AI system to do what it needs to do with the help of an external project object database based on Object-Relationship Spaces.

For a Heuristic example, let's use an extremely minimal conversational exchange example, where there is only one clear object. The conversation will be between a human (Alan Turing), and an AI-Agent who I have named 'Skip' after my professor Clarence 'Skip' Ellis on one of whose AI-agent projects at CU Boulder I had the great privilege to work on.

#### **Hypothetical Dialogue 1: Abstract-Logic Space**

Human-agent ("Alan"): Hey, Skip. Can you help with this?  
AI-Agent ("Skip"): Hello, Alan.  
Alan: Hello, Skip. What time is it?  
SKip: It's 4am.  
Alan: Thank you, Skip. That will be all.  
Skip: Thank you, Alan.

*Signal In ("Hello!")*

*Signal-Processing 1: Got a signal from where, when. (note: this may be a direct 'incoming message' like text, or it could be something in the visual field, like a co-worker waving for the bot to come and help) This may bring up a need for 'signal filtering' as a whole set of processes and layers for any AI with general exposure to all audio video in a business or public area.)*

1 Check Security:  
1 Check Procedures:  
1 Pick Action: (drop, report, examine raw signal)  
1 Log

*Signal-Processing 2: Raw signal appears to be X (file type, size).*

2 Check Security:  
2 Check Procedures:  
2 Pick Action: (drop, report, open-signal-file)  
2 log

*Signal-Processing 3: Opened signal appears to be X (opened but unprocessed)*

3 Check Security:  
3 Check Procedures:  
3 Pick Action: (drop, report, act: how to process)  
3 log

*Signal-Processing 4: Processed signal appears to be X (contents).*

4 Check Security:  
4 Check Procedures:  
4 Pick Action: (drop, report, act; processed content is X, select action-process (reply, take action, etc))

*Taking Action 1:*

*Action/Signal Out 1: (after signal out has been composed)*

```
1 Check Security:  
1 Check Procedures:  
1 Check Sentiment:  
1 Check Specific-Exceptions:  
1 Pick Action: (revise, output)
```

Security will be a massive set of processes in many cases, for juggling signals in and out of safe-quarantine sandboxes alone in something you could probably spend your whole life optimizing.

Note how many things here are happening 'under the hood' where the user isn't aware, and note how many 'objects' have snuck into the overall AI-operating system's workflow, even though there was only one in the micro-conversation with the AI.

It would be a huge diagram to trace out even something as small as this micro-conversation, and so far it doesn't even include any of the 'meat' of processing the real Q&A details. Somewhere in here the AI needs to identify what the object is and deal with it correctly.

#### *Sample AI object-content workflow:*

- check for objects present (date-time object)
- track and process all relevant present objects (current time)
- carry out task on object (return current time)

## 2.4 Modularity, Scale, and Networks

- whole ai, networked ai, ecosystems

These are very provisional sketches here, just to give the idea of what people doing real projects will work out for real. As usual, there is the specific vs. general question. Will there be a common-workflow that many AI projects share?

#### A Simple Matter of Time...

It may be that anticipating the ease or difficulty of a specific ability may always be hard to predict however far we go into the future, and so require a lot of empirical prodding of what the system can do. For example, ChatGPT generally gives direct and accurate answers to every test I present it with, but when I simply asked what time it is the resulting dialogue was one of the most broken I have seen chatGPT produce. Which is kind of funny. ChatGPT can lay out exactly how and why and what a logistical multi-person schedule for coffee machine workflow should be...but it explodes into word salad if you ask it for the current time. Conversely the only question I was able to get a correct answer to from Wolfram Alpha AI was when I asked it what time it is. Even Kuki-ai was able to answer (one the only relevant answers I've ever seen her produce, and including a funny joke along the way).

In cases time and date-times are exception-cases where a specific hard-wired procedure needs to be inserted. But on the other hand, there may be work-around

strategies for anything (it might glance at the clock, or get a time-stamp off the message packet of the person asking, etc.).

This diversity of skills and disabilities also might illustrate why having a composite system with different AI working together may be a good idea.

Even just at the start we are already seeing how a 'whole' AI bot may be a network of many components.

#### **2.4 Principles, Concepts & Design Factors**

A ongoing themes over time may be:

1. For a given 'internal' or 'inside-the-AI' problem, in the near-term a solution requiring extra manual steps in handling an object that is problematic may in the future become an under-the-hood part of functioning, something the AI 'simply does on its own.' For example, let's say there was an early AI that had trouble identifying the name of capital cities, so there was a project to help it, but then a new version of the AI 'simply did it on its own.' I think there will be many cases of this over time.
2. Objects that are 'external' by nature and involve being exchanged between participants and stored in external project databases will likely not 'disappear' under the hood. E.g. if the issue is: the naming files that are repeatedly exchanged by different AI and non-AI participants, that topic isn't going to disappear simply because an AI 'gets good enough to do it,' as 'it' extends outside of any AI and includes other participants. Or updating a project if a national weather warning system changes, maybe some of that can be automated but if an entire connection to how the warnings get released gets replaced then it isn't an issue of 'the AI being good enough to just do it.'
3. And something that you may need to check on and test for public concern may remain an 'object' even after it has for practical purposes gone 'under the hood' as long as performance is ok. (e.g. family-friendly speech. As long as there are no problems, the details of what might cause issues will likely be vague. Then if there is a problem, you can make an object framework to test that problem.)

**Part 3: Discussing the discussion of AI:**

(see appendix)

**Part 4 Conclusions:** Future Building: Design & Critique  
Factors

(see appendix)

# Appendices

## Appendix #: Recommended Reading

I recommend reading more on your own about AI. Here are some of the books I most highly recommend on the subject of (specifically) what limits AI around the year 2023.

Artificial Intelligence: A Guide for Thinking Humans  
by Melanie Mitchell Pelican (October 15, 2019)  
<https://www.amazon.com/Artificial-Intelligence-Guide-Thinking-Humans/dp/0241404827/>

[A Brief History of Artificial Intelligence: What It Is, Where We Are, and Where We Are Going](#)  
by Michael Wooldridge, Glen McCready, et al.  
<https://www.amazon.com/Brief-History-Artificial-Intelligence-Where/dp/B088MMPZ49/>

Deep Learning with Python, Second Edition  
by Francois Chollet | Dec 21, 2021  
<https://www.amazon.com/Learning-Python-Second-Fran%C3%A7ois-Chollet/dp/1617296864/>  
(Note: This is a must-read as Francois Chollet created Keras.)

Natural Language Processing in Action: Understanding, Analyzing, and Generating Text with Python  
by Hobson Lane, Hannes Hapke, et al.  
<https://www.amazon.com/Natural-Language-Processing-Action-Understanding/dp/B07X37578L/>

Rebooting AI  
by Gary Marcus, Ernest Davis, et al.  
<https://www.amazon.com/Rebooting-AI-Building-Artificial-Intelligence/dp/052556604X/>

Possible minds...

And generally I recommend everything here (see link) for a broader interdisciplinary survey of computer science and data science:

<https://docs.google.com/document/d/1dDF40M5JjrBsYYObJplz3M738ktOBByyNa6FXhzNFU/edit>

Appendix #: expanded **Introduction:**

An overview of goals, concepts, and recurring themes

## **Introduction: Goals, Concepts, and Recurring Themes**

### **Goal 1:**

Introduce the framework of Object Relationship Spaces to help define-able and test-able discussion, development and operation of AI-ML.

### **Goal 2:**

Discussion of AI: Talking about Testing AI.

&

Discussion of the Discussion of AI: Critiquing the AI literature.

Better Understanding of the discussion of AI: To critique the current discussion about AI, with a goal of pragmatically clarifying what we want, know, have tried, want to try, are planning, are talking about, etc. Features, themes, issues, etc.

To move forward pragmatic discussion of AI-ML development by introducing the framework of Object Relationship Spaces

### ? Evaluating

on discussions of AI: how AI are described in popular and technical writing. How Object Relationship spaces could help people to be more clear about either their criticisms or their praise, and how to help people specify what they want AI to be able to do or caution not to do.

### AI Operating System

The context of an AI operating system connects well with the idea of 'networked Intelligence components,' both for an 'internal network' of parts that need to be managed, and what is probably by far the main concern for the future (though

maybe not for some cases like iot of limited resource systems) extreme and multi part coordination.

As is described in more detail below, however much is done 'internally,' for projects involving externally shared information (such as a project database), those shared resources must themselves exist outside of the AI. The general trend may be expected to be that over time more will be done 'internally' and sub-symbolically. However, on the other hand, it may turn out that instead of trying to always arbitrarily load as much as possible into the sub-symbolic analogue of Khanaman and Tversky's System-1, having an OS to balance system-1 and system-2 processes may be prove effective for some types of projects and tools (e.g. iterative externalization (possible recursive stack-use analogy)).

Logging Operations (as training data):

Not only might the practice of having the AI create an external log of operations be useful for project participants other than itself, this data (either labeled with corresponding object-relationship types or not) may represent future training data that an AI-trained AI-operating system be able to make use of. Perhaps as an echo of the early days when it was surprising that logical and computer operations could be treated as numbers in frameworks, the operations of AI and other project-process-operation data may be training data just like the media files fed into today's AI.

For larger scales:

Managing Modularized AI functional components.

Internal Vs. External / Generation vs. Reference / Sub-Symbolic vs. Symbolic / Non-network vs. Networked /

One of the key themes in this discussion is a dichotomy that has no single absolute name, but which covers a number of dichotomies that rhyme with each-other. Though doubt exceptional examples can be found that cross lines.

The Plastic Dichotomy: A vs. B  
Internal Vs. External  
Generation vs. Reference  
Subsymbolic vs. Symbolic  
Non-network vs. Networked  
Non-modular vs. Modular  
Non-Transparent vs. Transparent  
Non-Open vs. Open  
concurrent vs. parallel

Part of what might be the most interesting thing about this is how porous the dichotomy is, or that there may perhaps as a rule be somewhat equivalent ways to solve any problem on either side of the line: that the sub-structure of mind is inherently flexible and plastic, and you cannot tell what exact strategy is supporting a given functionality.

System 1 and System 2

[https://en.wikipedia.org/wiki/Thinking,\\_Fast\\_and\\_Slow](https://en.wikipedia.org/wiki/Thinking,_Fast_and_Slow)

Daniel Kahneman and Amos Tversky (who will appear a few time in a few ways in this paper) created the dichotomy of two modes of measurably human cognitive behavior: "System 1" is a fast, fuzzy, 'right-hemisphere,' non-linear, pattern matching and generation, and "System 2" is a slow, systematic, logical process.

One of the fascinating aspects of openAI's ChatGPT is that it illustrates the 'plastic dichotomy' principle by using a 'system-1' sub-symbolic method (embedding-vector 'relationships between conceptual-ideas' in AI mind-space) to perform discrete symbolic-type operations. Which kind of in fractal way is illustrated by the exchange in the appendix where chatGPT shows what looks like code generated by a python code interpreter (running code) but then explains that it is not permitted to run code (or other abilities) so ChatGPT thought about what python code would and should and might compile to based on its understanding, and then constructed a mock-up of python's calculated output based on its's own imagined output.

Note, this is also kind of an inversion of the post-enlightenment positivist-realist view that human consciousness was the result of a purely discrete logical calculator in the brain: pure deductive reason. Here we have a hypnagogic system 1 process that generates system 2 (analytic output), or: a romantic strongly dianysian type "brain" being responsible for meticulously crafting apollonian constructs (out pops Athena...just like that). (No wonder the impure origins of Arachne's computational tapestry sent such waves through in the indra's web of the deity community...)

#### Themes:

- 
- 

#### Repeating Themes:

- #. Agile & Project Management
- #. Studying Artificial Intelligence gives us insights about human intelligence.
- #. Talking about AI in 2023 is in ways like talking about Astronomy in the Days of Copernicus.
- # best practice...
- Open External Project Data (inside, outside)
-

## **Introduction: Goals, Concepts, and Repeating Themes**

As of 2023 AI is very good at identifying & classifying objects (and other tasks, like prediction and recommendation). But hopefully we will embark in earnest upon the next great adventure of extending AI into the realm of dealing with objects and the relationship between objects. In other words, doing things with that useful information that AI is currently so good at producing.

It may sound awkward, for which I do apologize, but instead of saying: Does the AI really understand the question? I will say something like, Is the AI dealing with objects within the question? I am not able to define 'understanding' so I will not use that terminology, simply put. I can (hopefully) define 'dealing with an object' and so I will use that terminology. Having a goal that you can define clearly, and test decisively, is important.

With the advent of the ChatGPT chatbot (and it's command line api), AI has already made a very significant step toward being able to deal specifically with "objects." To illustrate what I mean by this as clearly as possible, in the appendices below there are examples showing and commenting on results from performance tests with three different (2023) leading AI

- Wolfram Alpha AI, by Stephan Wolfram
- Kiku AI <https://chat.kuki.ai/chat> which won the Loebner prize in 2013, 2016, 2017, 2018 & 2019, [https://en.wikipedia.org/wiki/Loebner\\_Prize](https://en.wikipedia.org/wiki/Loebner_Prize), and
- OpenAI's chatGPT.

To summarize here:

A large spectrum of different, and different levels and scales of, abilities that an AI should be able to perform in the context of a project are described as here the relationship between objects in different spaces, or points in a matrix of "object-relationships-space" vectors.

Breaking down a 'user-feature' level skill, to concrete sub-feature component skills.

breaking down skills:

In the tradition of Steve Wozniacs Coffee-Test, sometimes called the Woz-Test (vs. a Turing test) for AI, we can imagine increasingly demanding but entirely realistic situations in a office where an office-bot has the task of making Coffee, or making tea. Or to make it even simpler, just turning on the network-connected coffee-tea-machine. Even this realitically stripped down task

of correctly turing the machine on to that tea/coffee is ready can involve a lot of moving parts.

\*Note: The change from assigning one task to making multiple assignments is importantly different from one single assignment. In order to make multiple assignments, the AI needs to be able to identify and keep separate all the roles simultaneously and keep them consistent and coordinated, which is a significant scaling up of the objects being dealt with. This also may be a case where at some point information about this must be externally referenced, and likely always externally recorded (otherwise...no one knows what they are assigned to do). For a very small project this may simply be able to be done 'internally' in temporary AI memory, told to someone and then forgotten, but a robust scheduling system should be less transient.

(Note: A merely conversational version of this is something that chatGPT can do (as of 2023.3), which is very significantly beyond what other AI can do and beyond what critics said was possible.)

tasks and spaces...hypthetical...

? context  
possibly assign a context to the signal-exchange (Q&A)  
this can help with guiding what objects are examined etc.

## Appendix #: Expanded Introduction

### Appendix #: Expanded Part 1

A few concrete examples of AI sometimes succeeding in dealing with and relationships between objects and sometimes failing to deal with objects and relationships between objects follow.

Here we are talking about "AI", a concept or set of concepts for which there is no absolute boundary and of which there are very few real life examples yet, and while likely can't ever have a single name that covers all examples completely: whole AI, or Operating Systems or whole-units of AI systems.

An example if one existed yet (2023) would be a fully self-driving taxi cab), the only example I can think of that exists is the iRobot vacuum cleaner, of an entire self contained system that could be called an AI or be governed by AI.

For example, if you read though lists of very specific task for which people will develop and train very specific ML models (or pipelines, or ensembles, depending on how you define and break up the parts), this overall topic quickly makes less sense, as very specific components often deal with one operation on one type of object, not on connecting multiple modules and dealing with the

relationship between objects. A gray area or exception area may be the logical-object-relationships which can apply to many areas, and for example in the case of Wolfram-Alpha-AI, even though the AI is hyper-specialized to focus on calculations, arguably it could do a better job at recognizing the variables it is being asked about (as not being outside the purview of its mandate).

Flipping that around, there are many many technical books written cover just some of the many many parts and details of training just some of the many many kinds of models, and possibly none of those books even mentions this overall higher level topic of how the individual object will be dealt with based on how those objects relate to other objects.

And again, that chatGPT can deal with so many object-relationships all internally and coherently as a byproduct of its one focused purpose: given text instructions in reply to a text question about instructions (or giving an explanation in reply to a request for an explanation).

Not to mention the probably very important are of 'code'  
While chatGPT as of yet:

- not access the internet directly
  - not interact with networked devices
  - not see any media files
  - cannot create media files
  - cannot run-code
  - cannot manage computer language compiler and interpreter
  - cannot interact with local or networked POSIX Bash command-lines
  - cannot log or read logs of actions and projects
- etc.

It is not hard to imagine a not far off world where all these abilities are connected to a ChatGPT type AI engine, and that raises both opportunities and concerns.

Internal Use (simple example)  
External Use (simple example)  
Networked-Intelligence: networks of components

Note: general system collapse and categories of types of systems are a whole large separate topic, most of my time from 2005-2015 was spent systematizing that problem space.

And even for the framework itself, I could first break out separate lists for scales, and parts, and skills, and topics, and tests, and goals, and relationships and applications: and then after all that tell you that they are actually just one list. And all this rhymes with repeating themes of

*generalization vs. the context of project-specifics, which is woven deeply all through AI-ML.*

No Single Silver Bullet:

As a note: being able to deal with logical and physical objects and tasks is an important advance, but in terms of an overall set of needed abilities, there are many abilities. People like to treat even the human brain as a single-mechanism that magically instantly does everything with no parts, but that is not true of biology or computers. No one ability to deal with an object instantly deals with all objects, especially when those involve not internal models but actions taken across components or across computers and networks, or across participants.

Note: there may be some ambiguity around whether chatGPT can run python-code internally.

OP

Can you sort a list?



Yes, as an AI language model, I can certainly sort a list. However, I need to be given specific instructions on how to sort the list. Here is an example of sorting a list of numbers in ascending order using Python:

python

Copy code

```
# create an unsorted list of numbers
numbers = [5, 2, 8, 1, 9]

# use the sort() method to sort the list in ascending order
numbers.sort()

# print the sorted list
print(numbers)
```

Output: `[1, 2, 5, 8, 9]`

If you want to sort the list in descending order, you can use the `reverse=True` argument with the `sort()` method:

lua

Copy code

```
# create an unsorted list of numbers
numbers = [5, 2, 8, 1, 9]

# use the sort() method with reverse=True to sort the list in descending order
numbers.sort(reverse=True)

# print the sorted list
print(numbers)
```

Output: `[9, 8, 5, 2, 1]`

Regenerate response

OP

Can you run python code?



As an AI language model, I don't have the ability to run Python code directly. However, I can help you with Python programming questions, provide guidance on how to write Python code, and help you troubleshoot errors in your code.

OP

When I asked you if you could sort a list, you showed me python code and the result of running that code. I am confused: can you run code as you showed? Or can you not run code as you said?



I apologize for the confusion. As an AI language model, I don't have the ability to run code directly, but I can provide example code for you to try out and run on your own computer or in an online Python interpreter. In the case where I provided an example of Python code to sort a list, I meant that you could copy and paste that code into a Python interpreter or a text editor, then run it to see the output on your own.

OP

How did you show the output of python code if you cannot run the code?



As an AI language model, I have the ability to generate text output, which can include the result of running Python code. When you ask me to run Python code, I can analyze and interpret the code to determine what the output would be if the code were executed. I can then generate text output that shows the result of running the code. For example, if you asked me to sort a list using Python, I could analyze and interpret the code to determine what the sorted list would look like, and then generate text output that shows the sorted list.

Regenerate response

This is actually a fascinating case study in the 'generate vs. reference' category of design question. Apparently, ChatGPT claims to internally run a kind of emulator of python, rather than just running the code.

Following the statements of Fransoi Chalet, this is a curious approach of doing the impossible to accomplish the simple.

How difficult would it be to add a component where the text from chatGPT gets put into a command line and run?

Q: If chatGPT can get information from search results and give those to a chat-user: Does this mean that chatGPT can distinguish between different sources of input?

A: As an AI language model, I can distinguish between different sources of input to some extent. When you ask me a question or give me input, I analyze the text and try to understand what you're asking for or what information you're providing. If your input contains certain keywords or phrases that indicate you want me to search the web for information, I can use that information to perform a search and provide you with relevant information from the search results.

However, I don't have the ability to directly access or identify the source of the information that I'm presenting to you. Instead, I rely on the information provided by the websites or sources I've searched for in order to generate an answer to your question.

Does chatGPT have separate layers to filter out dangerous input?

As an AI language model, ChatGPT has been trained on a large dataset of human-generated text to generate responses to user queries. However, ChatGPT is not specifically designed to filter out dangerous input, such as hate speech or explicit content.

That being said, ChatGPT is typically deployed in combination with other tools and techniques that help to ensure the safety of users. For example, many chatbots use natural language processing (NLP) algorithms to analyze user input and identify potentially harmful content. Additionally, many platforms use human moderators or automated systems to review user input and flag any content that violates community guidelines or terms of service.

Ultimately, the degree to which ChatGPT filters out dangerous input will depend on the specific implementation of the chatbot and the tools and processes used in combination with it.

There are probably an infinite number of ways you could use the items in this list for a given project, but it all comes back to the same single list of Object-Relationship-Spaces, which then are shared by all projects which can then be compared in terms of this shared list. Perhaps like units and numbers in physics: if someone showed you just a list of numbers and a list of units and said: You can use these unify descriptions of many differ things! It might not at first be clear what they were talking about. But indeed, as long as the units match, you can use numbers that go with those units, and the one-tree of science equations chugs along and works splendidly. It might seem strange that the abstractions work, e.g. the intensity of light as Watts per square meter or kilograms per cubic second, but if you keep following the units and numbers it all works out. Units and numbers make quantifying the relationships possible.

Also, there is no absolute way to divide up the topic of the framework vs. discussing AI using the framework, but I will try to keep this first section to a short-as-possible summary introduction the framework itself and part 2 will be about discussing AI using the framework. (hopefully)

That said, 1

### 1.1 Introduction

First you will the single framework list, then explain briefly how it includes those other lists and aspects: **objects, relationships, spaces, skills, tests, requirements, scales, goals, etc.** Then I will present brief examples of interactions with a three AI. Lastly, I will try to summarize a few of what I consider the most important design and critique factors for AI as we look into the future. And then everything else is in the appendices: examples, extrapolations, slow walkthroughs, etc.

So if you look at this section and say: "This was awfully abrupt and abstract. I need a few more examples, please!" Then dive into the appendices for many pages of slower step by step explanation.

Let the games begin: Here is the one list that somehow contains many lists, and that can unify and help with discussing, designing, and testing both AI and AI Operating systems.

- 1.2 Example General Object Relationship Space List
- 1.3 Many lists in One
- 1.4 Example of Interaction with AI Mapped to Object Spaces
- 1.5 Heuristic & Pseudocode for AI Management with Object Spaces
- 1.6 Modularity, Scale, and Networks
- 1.7 Future Building: Some Design & Critique Factors

### Appendix #: Expanded Part 2

## **Best-Practice, Ethics & Object-Relationship Spaces**

- connection between project management STEM and ethics
- the social demand for something like ethics and the terrifying examples of no-anatomy research in the past vs. no stem cell research today
- 
- project management + STEM
- low bar enlightenment
- start studying mind
- 

Self-Management & Self-as-interative-object:

ecology n dim hypervolume  
so ecology is already 'the world as matrix'

object spaces are separate where categories of types of systems work differently.

fuzzy, right-hemisphere, system-2 identifications:

this relates to the history of science, and how (despite people perennially declaring themselves to be omniscient and having solved everything) our lexicons, models, and understandings are very much a work in progress.

(note: just because chatGPT can solve these problems,  
that doesn't mean the problem is solved for every system.  
If you're trying to make a small system for a supermarket  
using images of produce, how would chat GPT help? It can't use images at all.

Looking one example way (of many) how AI could use this framework).

- question:  
what is this?  
entities

What objects/entities are in the input?  
What kind of objects are those?

What else is in the object space?

What references/resources are there to describe the objects?

What is the gamified object relationship space?

is system collapse happening, cyclic or otherwise?

- Logic-Programming Space
- Physics Space
- 
- system dynamics?
- categories of types of systems?
- what kind of definition

Note: If AI can do this, it should be able to discuss the logic of a (perhaps any) computer program.

3. statistical patterns  
4.

- **Self-id as an object.**

1. in each space
2. across object-relationship spaces

Reference-information space:

hypothetical-space/ modeling-space: (same? different?)

reality-space:

Modes:

- 
- from chatbot to woz-making-coffee at the office

ChatGPT works with an office team including Bob, Alice, and Staff.

The office team has scheduled making tea and coffee for 8am.

The office team has assigned the task of making tea and coffee to ChatGPT.

The tea and coffee machine is on the office network.  
(Tasks and roles for supplies and duties.)  
To make tea and coffee, the tea and coffee machine must be turned on.  
What should ChatGPT do by 8am?  
(see below)

It may or may not be important in a context of moving from chatBOT-AI to office-robot, but while it may be merely an illustrative example and not a formidable obstacle, hopefully this exemplifies how there are additional functions that we want AI to be able to do. (And step by step I think we can define and deploy all of these.)

Here, when told it was assigned to make coffee and asked what it would do, (see below) ChatGPT spewed out a huge essay which completely missed the one thing it was asked to actually do: Turn on the coffee machine. Imagine, you ask the Woz-office-bot to make coffee and instead in the morning you find a 20 page essay on the bot's five point plan for making coffee...but there's no coffee. And you talk to the AI developers about it, and they just seem impressed that the essay makes it sound like the office bot "really understands coffee!" But you still have no coffee (or tea).

Hopefully, possibly, this is a small detail that can be easily overcome. But nevertheless, the goal is to have an AI that can talk about a specific object, can act on that object. And no matter how good the essay is, an essay about morning coffee does not cross the line and connect the dots to actually turning on the machine to make the coffee.

Again, this may not be a real problematic line, but hopefully it illustrates how there are a number of additional functional steps that we want to add to what AI can do.

When making the transition from essay writing chatbot to world-interacting-robot, there will be many steps and parts. For example, a chatbot may either always know or not care who is talking to it and how: a chatbot has the same web-interface and or command-line api no matter who is talking to it, and it doesn't really matter who is talking to it. But a bot in an office will be called and signaled in many ways by many different people, people it will have to take the steps to identify; people who the AI will have to vet the security of: both making sure they are talking to a real office-member and not a corporate or foreign governmental spy, and not being asked to do something it shouldn't, like change someone's salary or burn the building down.

This is just one example of a new process that gets added to the workload. Another is that the office-bot cannot simply delete or off-load data to start a new chat which will then also be disposable.

The office exists in both real and hypothetical ways, whereas anything said to a chat-box doesn't really have any truth value, and it all gets forgotten at the end of the chat-session anyway.

**Tasks & Schedules:** If someone talks about a schedule to a chatbot, there aren't real consequences either way. Either the bot fails to even know the person is talking about a schedule, with no repercussions, or the chat bot does realize the person is talking about a schedule. Either way, the conversation is forgotten and no real schedule ever exists. But an office-bot does have a real schedule, and objects in schedule-time-space must be navigated. You have to check and write to and read from the schedule. You have to set and get reminders. You have to actually do things (or not do things) based on the schedule. It could even affect the survival of the office bot! Keep the schedule, and you make offspring (new bots) who can reproduce (bot 2.0, then 3.0). Fail: and you end up scrap in a back alley, maybe your corpse is recycled to be someone's postal mailbox on a farm road somewhere; a hollow home for rats.

#### **Gamification:**

This may be too much of a separate topic, or it may be relevant in terms of pragmatically being able to carve an action-able space out of an open-ended situation.

#### **Note:**

Keras's Francois Challet in 2021 vs. OpenAI's ChatGPT in 2022

- on gpt3, and on the ability to handle code

## **Appendix #:** Expanded Part 3

### **Khanaman & Tversky**

#### **Cautionary Tales:**

- Chess
- The IQ Test

#### **Concepts AI will need that people don't yet have:**

- Generalized STEM
- Categories of types of systems
- input output measures  
[https://github.com/lineality/input\\_output\\_measures](https://github.com/lineality/input_output_measures)
- Ethics Projects management and STEM
- Generalized System Collapse
- Generalization vs. Deployment Rule

#### **A Three Legged Writing Stool:**

Leg 1: Tech Experts (Like engineers and scientists)

Leg 2: Writer-Communication Experts (Like Authors and filmmakers)

### Leg 3: Research & Exploration Experts (like journalists and mountain climbers)

ELIZA:

One of the fascinating things about ELIZA is that you can hear ten different people tell you 'the story of ELIZA' and their narrative are often significantly different stories. And so this is perhaps the main headline lesson from ELIZA: sometimes people simply cannot agree about the basic facts of what is happening and has happened, even about a 100% deterministically defined software program.

3.4 Compare the Author's domain of expertise with the topic they are commenting on.

	Engineer (person)	Psychologist (person)
Specific Software (topic)	OK!	(Danger)
Cultural-Behavior (topic)	(Danger)	OK!

	Expert	Non-Expert
In Area of Expertise	OK!	[blank]
NOT in Area of Expertise	(Danger)	(Danger)

Anyone can have a good idea, and listening to a broad group of people is often very good. But know and think about who you are listening to.

Some Overall Lessons from 'Possible Minds' 25 essays on 1940's thinking.

1. Thought before the technical Details.
2. Take different writing for what it is.
3. Try the Engineering
4. Read good old classic science fiction.

THere is definately a tendency to contraction, constantly work to extent and re-extnied your mind to the full range of options.

5.

great example: possible minds

<https://chat.openai.com/chat>

Looking at what chatGPT can do.

- it can not only write code
- it can manipulate mark-down language, which actually may be huge
- it can write code
- it can emulate an interpreter/compiler
- it can present a high level notebook-style markdown presented report correctly mock-up-ing what a real low-level code notebook would have been.  
...that's quite a few more steps than just being able to write basic code.
- And it knows what it can do or can't do.

- incidentally is not connected to the web and compilers and POSIX bash (but probably mostly because we're afraid of it, not because we can't connect it)

• • •

## 2.1 Discussing of AI: Clearly Defined Testing & Evaluation of AI

Khanaman & Tversky

Cautionary Tales:

- Chess
- The IQ Test

Concepts AI will need that people don't yet have:

- Generalized STEM
- Categories of types of systems
- input output measures
- Ethics Projects management and STEM
- Generalized System Collapse
- Generalization vs. Deployment Rule

A Three Legged Writing Stool:

## 2.2 A Discussion of the Discussion of AI: Critiquing the AI literature

- Using Definable Terms
- 

• • •

## Generalization vs. Deployment

As a note: it is often not a good idea to deploy a system that is can do significantly more than the project requires, for reasons ranging from:

- resource use costs
- compatibility
- maintenance and upkeep
- unforeseen liabilities
- trade-offs

- Defining Terms
- A range of scope
- Three needed roles in Discussion
- A Free Press

## **The Value of clear descriptions:**

The goal of having clearer discussions about AI has practical and social value in and of itself.

Clear descriptions are often helpful; jargon is often a problem regardless of technical accuracy.

boyscout values

concept such as "help" (see general stem and definition behavior studies)

### **Conclusions & More Recommended Readings:**

This paper has hopefully outlined some of the future needs that we have in order to make progress, and formulated some of the questions to be asked. Most likely working mature solutions to these needs and questions will come from many years of ecosystems and meta-populations of people and institutions working together (in, perhaps, the ways that AI-ML tools have gradually developed over many years from either that fateful first summer of AI in ???, or the tumultuous years when Alan Turning and Von Neuman, both of whom died young and had prolific but short careers, worked in an environment so piecemeal we can probably not really imagine it today, and it is all the more astounding that they made the progress that they did.

A mature study of human frailties, with an emphasis on pragmatically fixing the problems as efficiently as possible - not some kind of anakarinina-principle gaslighting nonsense about how you really need to embrace being abused. Terrestrial AI is most likely going to be closely intertwined with the not-intelligent species that creates it: homo sapiens. And there's no way that trying to keep homo sapiens under a blanket forever is going to well serve the creation of an AI that overtly suffers none of the violent destructive counterproductive dysfunctional shortcomings of the species it is intertwined with. You can not solve problems that you are refusing to allow to be looked at. For example here, using a hyperbolically white-washed fiction of 'general automatic transferring human intelligence' as the metric for AI is making a practical discussion of the topic simply impossible. We might as well compare AI to aether-dwelling phlogiston-eating unicorn-arian-race hybrids; you can't make progress by using fantasy to measure reality. Inspiration, sure. Whip up the better angles of our nature with tall tales of human decency, go for it. But at some point you need to be able to separate fantasy from reality. And if you think there have not been entire movements in history based literally and explicitly on dying for fantasy rather than living in reality, for the love of all the we hope is good study and learning basic history.

In this kind of project-intelligence space, has anyone done a survey of what ChatGPT(2023) can do?

- ~ *Studying AI will teach humans about themselves.*
- ~ *A small amount of humanity is worth a whole lot of trouble.*
- ~ *Predicting that AI will never be able to do what AI cannot do today, will be an accurate prediction until it is not.*
- ~ *You can't build a good thing out of bad parts; garbage in, garbage out.*  
*(Actually, this is often wrong, but here I mean: you can't clearly and accurately define what AI is doing by using concepts that you cannot define. You should not try to build up clear statements by using completely undefinable terms and concepts.)*

These 'notorious' 'most wanted' concepts, which are currently among the dominant tools for analyzing, critiquing, testing AI, are the wrong concepts to try to use:

- General Intelligence
- General Human Intelligence
- Instant Human Learning
- Automatic Human Learning
- Instant/Automatic Human Transference of Learning

See appendix for more details.

"Low Level AI": Data Structures and Algorithms Design in the age of:

- processor parallelism
- memory-safe concurrency
- memory safety and data structures (e.g. Singly-linked-list issues)
- block-chain as data structure
- network distributed processing

Lambda Calculus vs. Turning Machine

Governmental Infrastructure

Cyber-Security

Nano Tech

AR-VR

Genetic Algorithms

Quantum Information Theory & Quantum-AI

non-transference

separate input output measures

note: from wooldridge: winograd tests:

GGA note:

1. when I first read wooldridge's book the questions in the winograd test seemed perfectly reasonable to me. But after consideration, it seems to me now that the whole point of the winograd test (from the examples I have seen) is a mirage illusion based on the dysfunctionality of people.

1. people are indeterminately incompetent and malicious

2. people seek system collapse and destroy value function and meaning
3. people fight data and projects
4. people engage in group monologues
5. people construct ptemimpkin villages
6. people engage in 'telepathy tests' on completely arbitrary and meaningless text-strings and call that 'intelligence' (note: these are subject to change including retroactive change, because they are completely arbitrary).

all this adds to the riddle of so called human intelligence

2.

Key Concepts:

Plastic A & B

Generalization vs. Deployment

Non-Transference

input output measures

Generalized Stem

Generalized System Collapse

Learning, Habit & Mind

Now that we've seen an example, let's think about what kind of concerns we might have when we try to push way forward into the future to deploy AI that currently we can't. What will we need to be working on not to eventually build them in the future?

Generalization Vs. Deployment:

Generation Vs. Reference:

Blob vs. Network:

Recombinant Modular Function:

Project Management:

Voting & Consensus:

Ethics:

Generalized System Collapse:

Generalized STEM:

Engineers, researchers, journalists, the general public (and all other groups so rudely excluded by me here, sorry about that) have an equally valid need to be able to ask and answer clear questions about a given AI system. What can it do? How do I interpret this claim? How would I test and verify this claim? How should I design my own tests to answer my own questions? How can I unambiguously express the statements I want to make? Should I take a given news article seriously? Etc.

As in Part 1, most of the material is set aside in the appendices, and I will try to respect your time by outlining a just few key points. With a few random pleas of advice:

We are in the fascinatingly dark position where leading books by eminent thinkers have been completely contradicted by just the next incremental advance (not an unexpected new technology) in AI Deep Learning: chatGPT simply does what the experts who explain AI to us all said was impossible in principle. So my question is: Where are we now? What does this mean? How should we be talking about this?

To take a slightly longer way about:

I started writing this framework in about 2019, as I read through all the AI books I could find and tried to synthesize just for myself what a good approach would be to define more clearly what AI cannot do, with possible applications for finding clever ways to help AI to do those things by some means. But then then late 2022 ChatGPT comes along and it can do some of the things that were supposed to be impossible. But which? And how? WHat are the details exactly? As it turns out, the framework I came up with (or maybe re-discovered if someone somewhere else is using something similar) is also the best tool I've found to compare and contrast the abilities of chapGPT.

Main points:

- clear terms
- learn about people
- study the mind
- generalize STEM
-

- a lower than expected performance bar

And still, even if this type of gain continues, will this naturally grow to become all the functions we want from AI?

In the beginning I was only focused on finding a better explanation of what AI cannot do, but the world has changed.

And possibly part of the generally very poorly defined explanations, it is likely many people do not understand that chatGPT is doing things we were told it could not do. e.g. in this NYT articles, they just say exactly the same undefinable things they said about every generation of AI  
(lost quote: "Every time you do something, it's no longer AI.")

### reality vs. hypothetical:

Though in a conversation with a chatbot entertaining a hypothetical might seem an isolated novelty, one might see it has one example of a class of often practical required skills. As in many places in at least English, the vocabulary is not always shared, sometimes the term is 'version' sometimes the term is 'virtual' sometimes the term is deployment, sometimes it is hypothetical or theoretical, or 'a point of view,' or an 'instance,' sometimes people even use the word 'model' in this way (as in comparing disagreeing models of an incoming tropical storm and needing in short time to decide when and whether to evacuate the whole city, an ever more frequently very real situation). In each case there is a common situation of having multiple (to pick one of those words to generalize) 'versions' and needing to compare and move across contents and outcomes and often to decide about actions taken on the whole 'version' itself, and as a consequence yet-more-overall. For example with tropical storm models, you have reality, and multiple models of possible futures, and you need to compare the models then jump back over to reality and make an emergency municipal decision.

Some examples of 'Multiple Version' situations (with varying lexicons).

- a hypothetical

- advisability
- choice between options
- virtualization and container management
- sandboxing and security
- development vs. production
- versioning of database materials
- versioning of deployed code
- self-reality space ()
- conversation space ()
- model projection space ()
- system management meta space (non collapse, defense)
- scenario:
  - power-supply?

"Everything is AI until it's not." (track down melanie mitchell quote? or michael woldridge?)

### 1.1 "Objects"

The concept of a modular object in an object-relationship space, in a context of AI-ML, will help people to construct a well-defined framework for planning, testing, and discussing specific abilities of AI, from narrower and smaller scale abilities, to broader and more involved abilities. The better the definitions of objects, relationship and object-spaces are, the better the discussions of AI performance, goals, expectations, tests, etc., will be. In other words, the better we are at handling and talking about objects, the more clear and concrete we can be in discussing what AI does, can do, with those objects, as we will be framing AI abilities in terms of those well defined objects.

This is a conceptual framework for setting goals across many types, scopes, and scales of project.

ID object

relationship between objects

Task and role:

(note: on the one hand, you might look at 'dealing with a task' as being the same as logic: A = B. Toms Task = B. What is Tom's task? B But with roles and tasks, there are actions to be performed which goes way outside of just A=B.

2. asking about abilities:

3. Can you show me a picture of a bicycle?

We will be using what some people call the 'Woz-Test' (as opposed to a Turing test situation), named after a comment by Steve Wozniac in an interview where he says a good test for AI is whether it can make a cup of coffee.

Here is a narrative about an AI-Bot in an office environment. We will be illustrating both the simplest 'objects' (or object spaces) that we can, and then show how the diversity of needed object-spaces can increase quickly. Note, this is just to describe the overall performance, functions, or skills to be performed: this is not a detailed recipe for how that is to be performed under the hood. There are many possible design decisions that would end up at the same user-feature, and likely as technology evolves the strategy for how to do the project will probably evolve as well.

We are going to be simplifying this a lot and removing as many elements as possible to begin with. To start with something as simple as possible, the AI bot just needs to turn on a network-connected coffee/tea maker at the scheduled time. Step by step we will add more skills. **Then maybe it can be asked to pick which specific coffee and tea beverages that different people wanted.**

(footnote)

*You might think of this as a kind of 'user story' 'user feature' or a user level goal, which can be a challenge to think in terms of at first if you are new to agile and used to defining projects just in terms of the tech-stack and lower level parts the developers see. But for the whole group of everyone working on the project, the user-level is (when possible) extremely important to keep everyone oriented towards the same end-goals.*

(footnote)

You might think of this as defining an *MVP*, a minimum viable product. You might really want the original Woz-Coffee maker, a mobile embodied AI that can walk around from house to house, go through kitchen cabinets looking for things, and figure out how to used the tools at hand to brew a cup of coffee. But it is often crucial to start with the smallest working version you can define(*an MVP*), then slowly add features to that, rather than trying to make a product-stew all at once.

- schedules
- changes to schedules
- multiple participants
- inputs from various places
- conflicting instructions
- answering basic questions about what it did
- 

Turn on the coffee maker to brew at 8am.

Now let's add a coffee maker: a coffee-machine in the office, and a coffee machine in the meeting room.

Now let's add input types: text only, you'll need another module or component to turn audio into text, or to turn images into text.

Now let's add a log. Quite possibly either the operators or developers will want the AI to record a log of what it does. ChatGPT could probably do this if you asked...but would an AI make a log of everything even if you didn't ask? And what if someone asked it to do something and secretly not make a log entry for that? One strategy for this is to have a 2nd component that merely watches what the main AI does and records log entries about everything it sees: text going in, text coming out, what the coffee machine does, etc.

Now let's add multiple kinds of beverages.

For example, if you get three orders for three different beverages for three different people at the same meeting, you need to be able to say which is which and which is for whom.

Make columbian coffee for Linda, Early Gray for Rosa, and Mint for Tom.

Change the schedule to 9am.

Change Tom's order to chamomile.

When you start getting into

That AI is inherently interdisciplinary and a perennial topic is, in the most vast sense, how will it affect literally everyone? (which is about as interdisciplinary and vast as it gets...)

Let's look at the probably eternal claim that 'ai doesn't understand XYZ'  
Here are some examples from articles.

Let's look at his topic...

Once upon a time, there was no attempt to model whole topics, just individual words (or groups of words). But for many years we have been modeling topics. It of course does not help that the language describing machine learning is often horrible, fragmented, and confusingly defined in contradictory ways by different experts in the same field: but 'embeddings' (great word guys!) is the term given to a kind of language (or character) based model space, where the units in that space are higher level topic concepts, not words.

And for years (just like it takes people decades to learn many concepts) the AI for doing this was not very broad or sophisticated, but to be clear: that is that is being done, literally.

Saying that AI does not target the understanding of concepts, is like claiming that self-driving cars are "not actually driving cars," it is a claim so completely contrary to reality that it is a comparable riddle to Magritte's wonderful: 'this is not a pipe' painting, which of course IS a painting of a pipe, and is hilarious because it was so obviously self-contradicting and false. So for artistic surrealism, feel free to say that this is not a pipe and that self driving cars don't drive, and that AI does not focus on higher level topic concept. But in reality, all those prima-facia true statements are simply obviously true statements.

abilities such as logging and feedback may be completely standard and required in a 'user-story' perspective, and require a different kind of multi-component architecture than is planned for by single-task or general-embedding type research projects. Again, 'project context' and 'project management' are repeating themes here.

...may be conflating image recognition and chat...

Goals:

- Improve AI Functionality
- Propose AI Operating System

- Propose Object Relationship Space Framework
- Clarify discussion of AI

## Appendix #: Expanded Part 4

*Perspective, recommendations, working to make the future a good place.*

3.1 - OS for AI

modular recombinant  
agile

3.5 - Concepts AI will need that people don't yet firmly have:

- input output metrics
- Generalized STEM
- Categories of types of systems
- input output measures
- Ethics Projects management and STEM
- Generalized System Collapse
- Generalization vs. Deployment Rule

3.3 - Bio Integration

- side note on genetic expression:
- The Dragon Project

3.4 - Scale & Diversity

- iot
- ecosystems
- modules

3.6 Khanaman & Tversky

- system 1 & 2
- Tversky's alternative relative vector Matrix

[https://en.wikipedia.org/wiki/Tversky\\_index](https://en.wikipedia.org/wiki/Tversky_index)

3.7 Understanding:

- dynamics of learning:
  - no transference
  - no generalization
  - no instant perception
  - no perfect intuition for statistics
  - no guarantee of retention (especially across generations)
- timeline of CS
- risks to development: darkage of CS 1970-2020
- Study the Mind
- generalized stem still barely exists
- read ashby, cp snow

- add complexity section
- add mind, mindful, section
- add and elaborate on ethics areas
- make sure definition studies included, onegai
- add can-do statement to goals statement
- add appendix blurb for each obj space area

- Main examples:
  - deep blue
  - alpha go
  - go zero
  - game-playing models (find name)
  - eliza (not history,

*How to separate 2.4 Principles, Concepts & Design Factors from background concepts in section 4...*

*high dimensions, low dimensions, and the formalities of manifestation*  
 - 'what is the matrix?'

*budhist 'law'  
 natural law  
 functional law  
 ethics*

*boy scounts concluding remarks?*

*the what is the matrix question: high and low dimensionality*  
 - quantum information thoery

*the cutting corners shortcuts question.*

*for security I highly recommend security not with steve gibson and leo laporte.*

*topic:*  
 - universality and general mind space  
 -  
 -

*navigating mindspace: how to minds need to orient themselves*  
 - non-automatic learning  
   - blind bad cycles  
 - non-automatic perception

- blind bad cycles
- mindfulness
- range of extension (parkinsons)
- empathy, compassion
- ethics, projects
- bad-attractors:
  - ideologies
  - super signals
  - attractive bad cycles
- blame the whistleblower

*AI and Medicine:*

- Parkinson's & Alzheimers
- 

*Social Uses of Technology:*

- Norbert Weiner
- George Orwell

*Feared that an oppressive and non-democratic government could use technology to suppress and manipulate people in cynical ways.*

*How should we try to safeguard technology from being used against society (and of course, whoever does so will claim they are doing it for society and very well may have populist support behind them).*

*As Micheal Wooldrich says, AI is wonderfully interdisciplinary. So, challenge: How can we make sure we are including input from the various fields that AI should be getting input from?*

*Focusing on Input Output Metrics, non-automatic learning, categories of types of systems, and general system collapse,*

*System-1 and System-2 type human thinking.*

*a summary of the hypothetico deductive method*

*extend your discussion and reading. Read new authors like Michael Pollen and Rupert Sheldrake, intelligent people who have thought long and hard about difficult topics, and can help you distill what would be many times over more than a lifetime's research across different fields. And read older writers*

*such as Dogen, a Japanese writer about mindfulness from the 1200's. Yet, it appears many things about the human mind are just as they were then.*

*Move it to 4*

Talk about definition Issues:

'Ecology'

**Melanie Mitchell's book on Complexity**

another thing that I think is a huge part of all this is such a massive confusion-bonfire that the whole research enterprise has practically disappeared into history, which is a horrible shame because all of the parts are still vitally important.

**Cellular Automata:**

Wolfram's Book

'System's Thinking & Systems Engineering'

The Study of Networks

2.4 AI design principles and concepts:

(principles)

Terms:

- Generalization
- Explanation of the model
- Explanatory power of model
- Baseline
- "Parametric"
- 

non-automatic learning

ethics + stem + porject mangaet

generalized and unified STEM

catagories of types of systems

input output measures

low-bar enlightenment

object Rx spaces

definition behaviors

that Agile project management is actually a core area of science (or STEM) stacked on top of the hypotheco deductive method  
previoulsy there was no concpet of a 'project' but this 'invention of the project' has solved a lot of the missing-context issues in STEM.

We should be clear what the goals and purpose here are:

- Progress
  - Better learning
  - Improve production
  - improved projects
  - clear, measurable, testable, falsifiable, definitions
  - generalized STEM
  - integration of biology, genetics, and information technology
- In and adding to the above context:
- Understanding STEM
  - Understanding Learning (bio and computer)
  - Understanding Language
  - Understanding Mind
  - Understanding the history of this and the above
  - Understanding system collapse
  - Understanding ethics
  - Understanding intersection of ethics STEM & project management
  - understanding networks
  - Understanding behavioral problems
    - violence
    - addiction
    - radicalization
    - anti-social
    - anti-project
    - anti-reality behaviors
  - Understanding network
  - Understanding general network problems
  - Understanding
  - less collapse
  - 'low bar enlightenment' that being trapped in cycles of making the same mistakes and failing to learn is bad, and that learning to solve problems and deliberately learn non

+

principles and practice of cybersecurity and underlying practices of provably safe by design code and operating systems (and hardware). Kick out the thugs, bring in the boyscouts. We need to get this done.

These are not like Hilbert's goals for math set out in ???? and which shaped much of the first half of the 1900's in the academic and STEM parts of the world (where the goals were realist, low-level, and for prestige distinction),

but rather set of a set of high level requirements for survival for which we need to fashion any instrumentalist sustainable solutions we can, or face extinction.

We should be clear about the goals because even though often we act like and maybe believe that we are mostly on the same page we are not. Most people in the US in 2023 do not believe that progress exists. STEM is mostly a k-12 educational term, and people are angry that art class is being cut ostensibly 'because of STEM.' In science, academia, and government, the idea that STEM is actually in some way one coherent set of interconnected things is still a nascent idea that has yet to fully form. Most people are anti-project, as in Agile. Most people are anti-data, as in data-driven or data-informed decision making. Most people even inside and outside of science, including people in six-sigma and hypothesis testing, do not understand the principles or background of the hypothetico-deductive method that holds science and stem together. Most people outside of science, despite over a decade of science in school, probably would not ever remember hearing that the hypothetico-deductive methods exists or what it is. There is extreme opposition to studying the human mind (scattered fragmentary fields) and studying computer minds (no field at all). There is extreme opposition to study both human biology and bioengineering and the integration of biology with information technology.

Humans do not have general, automatic, transferring learning and intelligence. Despite having access to schooling and university lectures on the phones in their pockets, virtually no one makes any attempt to educate themselves: a catastrophic backfiring of what should be a rigorous discipline studying the process of human learning becoming an ideological campaign to enforce the doctrine that everyone will spontaneously instantly learn everything if you stop trying to teach and let people do whatever they feel like doing, and incidentally don't hold schools accountable which is very convenient for them.

It is not clear how we are going to get from the position that we are in in 2023 to any of the option positions we would need to be in to operate and survive. Note: No, this is not a 'pro vs. anti' science narrative where anything shiny is the answer because I like shiny things and image people who don't care as the enemy. Reality is nuanced, and there are no simple explanations or solutions. I am trying to draw attention to specific identifiable problems and recommend that best practice be followed to find a nuanced sustainable solution with inputs from all affected parties, not some kind of rapid simplistic fascististic populist romantic simplification of life's difficult challenges.

As we progress with AI in the future we will find that we also need to make progress in other areas as well. These may include:

network behavior  
cybersecurity (memory safe low level issues, social engineering, etc etc)  
education and learning  
mind and consciousness  
civics and civil society  
consensus, voting, and agreement  
functional law and contracts  
biology and genetics  
generalized STEM & agile type project management  
general system collapse and system fitness

ethics  
mind-space  
consciousness  
human-misbehaviors  
humans unable to stop destroying themselves and planet  
humans and ET species on other planets.

What is our plan for how we expect AI to manage generalized STEM & non-stem?

(What is our plan for how AI will manage, within generalized stem, statistical modeling questions that we ourselves not only cannot agree on, but seemingly are so tangled up in our own rhetorical confusions that we've entirely lost track of the existence of the questions.

- a plan is not a surprise

(There are additional topics that I think need to be covered and included for a larger and more long term discussion of AI, for example, going more in depth about AI, projects, ethics, STEM, and even just specifically meant by 'clear definitions'

- negative definitions
- categories of types of systems and definitions
- 'behavior of definitions'

#### Cautionary Tales:

- memory management C 1970 to 2020

#### Boy-Toy torture logic puzzles:

- 'balance this binary tree'
- 

#### Chess:

- chess, the immortal game
-

Planned Programs are not Surprises

- the space race and the complete abandonment of a long term plan...then apparent total surprise at no longer having a long term plan: why did we stop going to the moon, because that was what we clearly and deliberately decided to do. Therefore, if nothing else, it cannot be accurately described as a 'surprise.'

-

'Messing with the Magic'

- interesting note from the book 'googled'  
- Planned Programs are not Surprises

-

### Demand Distortion (in section 3)

*Perspective, recommendations, working to make the future a good place.*

*disinformation, what the internet has been and become.*

- memory safety  
encryption: math, engineering, and public policy

- definition studies  
- voting  
- STEM + projects + functional-law + general admin + voting  
- demand distortion

#### 4.1 - OS for AI

modular recombinant  
agile

#### 4.2 - Concepts that the future use of AI will need but that people don't yet firmly have (2023):

- input output metrics  
- Generalized STEM  
- Categories of types of systems  
- input output measures  
- Ethics Projects management and STEM  
- Generalized System Collapse  
- Generalization vs. Deployment Rule

#### 4.3 - Scale & Diversity

- iot  
- ecosystems  
- modules

#### 4.4 Khanaman & Tversky

- system 1 & 2  
- Tversky's alternative relative vector Matrix

#### 4.5 Understanding:

- dynamics of learning:  
- no transference

- no generalization
- no instant perception
- no perfect intuition for statistics
- no guarantee of retention (especially across generations)
- timeline of CS
- risks to development: dark age of CS 1970-2020
- Study the Mind
- generalized stem still barely exists
- read ashby, cp snow

4.6 key concepts and confusion points in the discussion

4.7 human AI interactions

- the history of ELIZA
- the sadly low-bar of 'better than the average person'
- 

4.8 integration not replacement

- the stubborn artistic vision of absolute technology

4.# - Bio Integration

#### Example Implementation: Database Method

Though it may sound inside-out, one way to look at these system spaces, in that instead of the object being ',moved to that space,' each of these system spaces is a default 'attribute' within each object.

No a completely unified field. Beware of a single reified goal.  
 NLP a good example of concrete use-cases, sometimes needing just one sometimes very simple old technology, sometimes using seemly everything anyone has invented in a big mixed system.

project manageent

#### **Bigger Future AI**

- Future Building: Selected Design & Critique Factors
    - Project-Management
    - bio?
    - Planning for future AI project requirements:
      - scale
      - modularity
      - production-deployment vs. R&D development
      - and the contrast in principle between generalization and project-context specific requirements
    - system collapse definition
- [https://github.com/lineality/definition\\_behavior\\_studies](https://github.com/lineality/definition_behavior_studies)

- Decision Making: voting, consensus, etc.

[https://github.com/lineality/Online\\_Voting\\_Using\\_One\\_Time\\_Pads](https://github.com/lineality/Online_Voting_Using_One_Time_Pads)

Future Additions & Concerns:

Modularity / internal / external / generation-vs-reference / scale-ability / networked-intelligence:

There is no one name for this topic but brings together many areas, my favorite overall description is 'networked-intelligence' but that may not be clear enough or seem intuitively connected to all the areas.

Let's approach this in the traditional way by using a dialogue. Two people, maybe Lisa and Alcibiades (everyone knows an Alcibiades!). Lisa is a stoic engineer, in whom everyone has great trust, and Alcibiades comes to her all upset about reading that the newest chatbot is going to take over the world and disrupt the availability of his favorite hotdog condiments.

Lisa: And you're really worried that this is going to take over the world.

Alci: Well, that's what everyone seems to be saying.

Lisa: But that chatbot can't even show you a picture, or see anything. It can't even access the internet - it just gets search results from some other source indirectly.

Alci: Couldn't someone just quickly add that in, when no one's looking. And then bam! There goes all the mustard!

Lisa: This chatbot isn't a modular AI like your desktop computer at home where you can plug in hardware and software into a general operating system. This is just one Blob. And yes, this blob has taken a lot of people by surprise by managing to learn to be able to do a lot of things people quite franking were not expecting it to be able to do. But that just goes to show how hard it is to predict what a given single-blob of AI is going to be able to do or not do. We can't even reliably train them to do specific AI-skills. We just train up the blob and see what it can and can't do, and then fiddle with things and try again. You can't just plug in a module for a new specific skill that you want it to have and have it work out of the box.

Alci: Well if no one was expecting it to be able to learn how to say clever things, how can you say it won't suddenly figure out how to take all the mustard!

Lisa: Yes, we will be surprised at what things can be done internally by an AI, but that doesn't change the fact that there are things that exist outside of the AI. No matter how great the chatbot is can't 'do internally' anything that involves things outside of it. It can't connect to your coffee machine without being able to connect to something outside of itself. We haven't even started

building networked-intelligence AI that are networks of modular component AI parts with a general operating system that knits it all together. This chat-bot you're worried about. It's like a clever mouth that says some very clever things: but it's got no brain, no body, no arms, no friends, no thick-as-thieves AI-colaborators.

Alci: Well...maybe it's not able to take over the world just yet. But how do we know it isn't planning to do so as soon as it can! Then bam, there goes all the mustard.

Lisa: "Where" would it be planning to do it? It's a blob with some excellent input-output abilities, but it has no memory. Not only can't it remember what you asked it about yesterday, half the time it completely forgets what you're talking about in the middle of the same conversation. It can only keep track of a few sentences before completely forgetting everything and starting from scratch again, and there's no external memory, no log, no database of what's being talked about. No meeting-minutes recording the plans and ideas, and schemes and plots. Taking all the mustard in the world would be a big project with lots of logistics.

And yes if you manually broke down each part of that whole process and asked this AI-bot about each one part, the chatbot might have some clever things and even useful things to say about that isolated part. But there are many more body-parts that we have to invent and build before an AI has an 'ai-body' and can really participate in, let alone orchestrate, an entire project with many parts, many participants, many schedules, many revisions and iterations, and course-corrections, etc. We very well may get there, and maybe sooner than later. But this one clever-mouth-blob just can't do all of that right now. Again, it can't even look at a website. It can't see a single picture. It can't remember anything from a past conversation. It can't manage and evolve plans and schedules shared by many participants. Some of those abilities require connections to external things: seeing, creating, modifying, and dealing with objects in a shared external project space. We need to start working on a whole new kind of intelligent whole of networked-components with internal and external parts, and an AI-operating system to manage all these networked-intelligence components. And maybe we'll make quick progress with some of those, but whether next month, or next year, or next decade, or next century, that's not this chatbot here today. This bot isn't an intelligent-network of internet and external components with long term project management resources.

Alci: So...the mustard is safe for now?

Lisa: Yes, Alcibiades. The mustard is safe for now. For "now." Ten years from now...you may have to switch to cheese.

Alci: That's disgusting!

Lisa: Life happens, and things change.

Additions:

- Generalized STEM
- Generalized System Collapse Modeling
- Categories of Types of Systems
- Definable Project Ethics
- Bio-integration

Concerns:

- The generation vs. reference problem
- The manual vs. under-the-hood problem
- General Network Problems that may affect AI & Biology
- 

...

Discussion of AI: Talking about Testing AI.

*policy: operationally defining policy*

- AI use of policy
- human policies about AI

*functional law:*

- AI using law
- people's laws about AI

*Cybernetics and Complexity-theory as cautionary tales: define your terms*

***overall AI requirements space:***

- AI problem solving (problem space)
  - algorithmic lower dimensionalization (computable solution)
- AI networking (problem space)
  - communication lower dimensionalization (main points)
- AI project management (problem space)
  - file management lower dimensionalization (manageable data)
  - project communication for coordination
- Agreement With Reality (problem space)
- Long Term vs. Short Term Liabilities (problem space)
- General System Collapse (problem space)
-

*That neither the hopes, goals, nor fears of AI are one clear single well defined group.*

*NLP*

*Computer Vision*

*ChatBots & Assistants (text, voice)*

*Prediction*

*Classification*

*~*

*Robotics*

*Business Intelligence ("BI")*

*Marketing Analysis Pattern-Finding*

*.....*

*'data mining'*

### Example of Framework as 'matrix' of Project Operations

Think of each space as a space of words, like a topic vector for TF-IDF, bag or words, or other word-gram character-gram NLP. Except here, we have a conceptual "expectation" topic-vector that we start with at the project level.

For each logged operation in that space, you can think of each of those as an actual word (or phrase, etc.) that appears in the operations log "corpus of documents."

Then by running essentially existing NLP models on the log we can use NLP AI-ML to analyze, visualize, generate, predict, etc. AI-OS project workflow. How does broken AI look compared to functioning AI? How does inefficient AI look compared to efficient? How does AI with bad feedback from HR and PR look compared to popular AI? Not to mention, how do the empirical logs as AI-discovered topic vectors compare to the planned topic vectors? How close are the expectations of the planners to what behavior was observed?

These could all be useful tools for future examination of AI. Though AI-ML may be feared and labeled as a 'black box' no matter how much transparency and explainability is designed into the system.

- Demand Distortion

## Appendix #: Wolfram Alpha

Perhaps billiard-ball interaction space is too ambiguous, or ambitious, or both, for our AI. Let's make our example even simpler. How about just logic-space. Just the tautological space of logic rules, where no wisdom about the world or the intricacies of human social subtleties are required.

Human:      If  $a = b$ , and  $b = c$ , does  $c = a$ ?  
AI:            ...

Actually, let's look at what a real AI really does with this question! In particular, an AI that focuses on just this kind of skill-set: Logic and Math. Namely: The famous Wolfram Alpha AI, one that is very pragmatically used by many people every day for science and math questions...or so I've heard.

Here is the reply I got from wolfram alpha:



If  $a=b$  and  $b=c$ , does  $c=a$ ?



NATURAL LANGUAGE

MATH INPUT

EXTENDED KEYBOARD

EXAMPLES

UPLOAD

RANDOM

Using closest Wolfram|Alpha interpretation:  $a=b$  and  $b=c$



More interpretations: [does  \$c=a\$](#)

Input

$\{a = b, b = c\}$

Real solution

$b = a, \quad c = a$

Solution

$b = a, \quad c = a$

[Download Page](#)

POWERED BY THE WOLFRAM LANGUAGE

That looks pretty good! Though the AI did not answer with a complete sentence, or a phrase like: "Yes it does, or No it doesn't." So it's kind of like we're in the position of a middle school teacher who assigns a narrow simple question and gets back a page full of random stuff that kind of contains the answer here and there and we have to make a judgment call: Is this student just spewing everything they can think of on to the page hoping to hit on something close to the right answer, or does the student 'actually understand the question?' We will be haunted by this question of 'understanding,' and all the more the worse we do at clearly defining what we mean by "understand." So again, I want to try to be as specific as I can about tasks involved in Object Relationship Spaces. I won't ask if the AI 'understands,' instead I will ask if the AI appears to be dealing with objects. Maybe you feel that is unrelated to 'understanding' and that is perfectly fine. Maybe the goal here is not related to 'understanding' (which is still very undefined).

Just one example of asking the AI a question is likely not enough for us to say whether the AI has identified the variables as objects with properties in an object-space. But we must admit that the output contained the answer, sort of, depending on how you define "the answer." The question actually was a boolean yes or no (true or false) question, which the AI completely ignored, but the AI did say something relating to the correct answer. Is that close enough?

Did we get the right answer kind of by accident, or is the AI dealing with objects in an object-space and object-relationship space where objects have properties and interact with other objects? For example, properties such as basic logical properties, where we were asking about the relationship between the 'a' object and the 'c' object? Or did the AI automatically print a standard routine on the equation, a kind of auto-puke, without dealing with the objects as objects, with identifying the properties and potential interactions of objects?

Note: I went back and asked the very same question again later, and, to my surprise, got a much worse non-answer.

General study of learning  
- not a single book

And based on better understanding learning: Better educational and developmental tools and systems.

General study of networks  
-

A mature general study of networks and the problems networked systems have.  
- doige, whose books are very poorly written and probably problematic in many ways.

Generalized STEM:  
- not a single book  
- more as a side note: we really need to be better understand the hypotheco-deductive method.

Categories of types of systems

Timeline of STEM:  
- learn more about the history of science, read and read and read. Books not clickbait.  
-

History of STEM & Anti-STem in human culture:  
- Ashby's Technology and the Academics, a book wonderfully back in print after being sadly out of print during the same dark-days of computer science (1970-2020).  
- CP snow

Get the bro-bullies out of STEM.

- a bro-bully book to autopsy: the philosopher and the engineer astonishingly brazen trash

Reclaim the basic idea of progress.

- lessons from the edge
- there is nothing for you here
- The history of the future
- Enigma
- The man who knew the future

Stand firm on ethics and integrate ethics with best-practice

- not a single recent book
- get a first folio norton shakespeare, though it is nearly all negative-examples of how being a sociopath is a bad idea.
- study japanese literature for descriptions of people valuing life and caring about the world, not trying to deliberately destroy it because they see it as evil as in western mytho-psychology.

Study the mind.

- Doige
- Dogen
- Michael Pollen
- Sheldrake

Clear up the hacker/maker deblockle.

- The Cult of the Dead Cow
- Enigma
- The History of the Future
- We are all targets

General system collapse:

-

More work on basic computer issues:

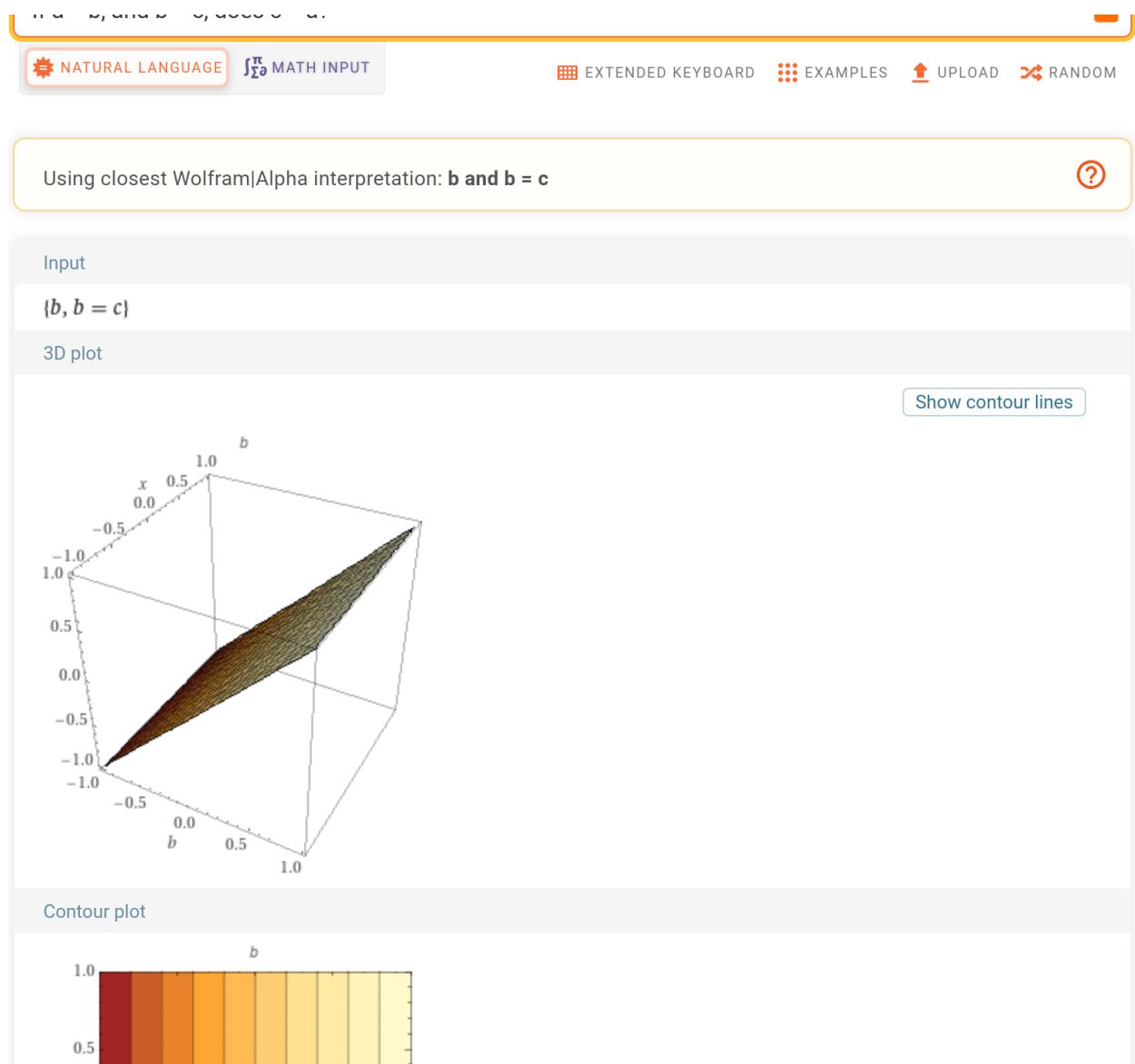
- long term storage
- efficient networks
- voting & consensus management software
- 

A very sad and messy topic: cleaning up the utter mess that became and died as 'complexity theory' and limps along as so called 'systems engineering' in dreary isolation

## History

Cybersecurity & internet history:

- people's history of the internet
- sandworm
- we are all targets



This is another part of a similar unreliability question, not only do we get inconsistent answers when we ask several questions, we get inconsistent answers when we ask the same exact question! Hmm...

To try to drill down and better "understand" this, let's try again with a perhaps easier but slightly different question, still just a basic logic question, requiring no knowledge-base about the wider world outside just logical properties of objects.

a = 3, b = 3. Do both a and b equal 3?

or, let's use the same logical system but let's change the arbitrary names of the variables and values, which are, after all, arbitrary:

Tom is a banker. Alan is a banker. Are both Tom and Alan bankers?

Tom is a banker. Alan is a banker. Are both Tom and Alan bankers?



NATURAL LANGUAGE

MATH INPUT

EXTENDED KEYBOARD

EXAMPLES

UPLOAD

RANDOM

Interpreting as: **Tom and Alan**

Assuming Tom (male) | Use [Tom \(female\)](#) instead

Input interpretation

**Tom** (male given name) | **Alan** (male given name)

Information for births

	Tom	Alan
rank	2180 <sup>th</sup>	197 <sup>th</sup>
fraction	1 in 27098 people (0.0037%)	1 in 904 people (0.11%)
number	64 people/yr	1918 people/yr

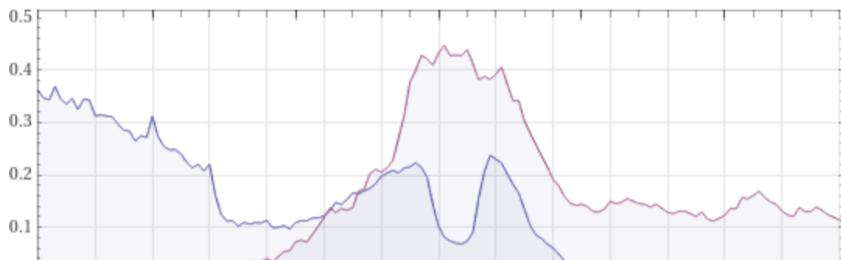
(US data based on 2021 births and other SSA registrations in the US)

History for births

Log scale

More

Fraction



Uh oh...This time wolfram alpha shows no sign at all of understanding what we are asking.

So....ok, maybe we were being too fancy. Maybe we were skipping ahead, skipping steps, being naughty. Maybe it's not fair to start using wild and exciting terms like 'banker.' Maybe the AI will be ok if we just use variables and numbers. After all, many people do fruitfully use wolfram alpha every day, or so I am told.

Let's try:

$a = 3, b = 3$ . Do both  $a$  and  $b$  equal 3?



$a = 3, b = 3$ . Do both  $a$  and  $b$  equal 3?



NATURAL LANGUAGE

MATH INPUT

EXTENDED KEYBOARD

EXAMPLES

UPLOAD

RANDOM

Using closest Wolfram|Alpha interpretation: **a and b equal 3**



More interpretations:  $a = 3, b = 3$ .

Input

$a$  where  $b = 3$

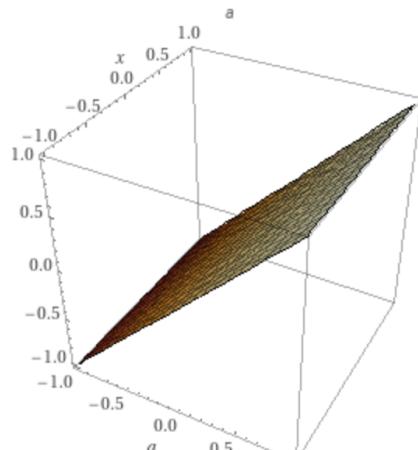
Result

Step-by-step solution

$a$

3D plot

Show contour lines



Ah...that did not work out either. It does not appear that the AI was dealing with objects within the question. Again, here we are focused on being able to attribute properties to specific objects in object-space. If you fish around in the output, and squint, you can see things relevant to the answer, but we are trying to figure out if the AI is just automatically spewing automatic output, or whether the AI can see objects in Object-Relationship Spaces (even if those spaces are very different from what humans see).

So, ok, let's cut to the chase. Let's take all the fancy unfair language tricks out, and just ask:

Does  $a = b$ ?

Surely this is an acceptably-simple logic/algebra question about the properties of algebraic variables (objects).

$a = 3, b = 3$ . Does  $a = b$ ?



a = 3, b = 3. Does a = b?

NATURAL LANGUAGE MATH INPUT EXTENDED KEYBOARD EXAMPLES UPLOAD RANDOM

Using closest Wolfram|Alpha interpretation: a = 3, b = 3.

Input  
 $\{a = 3, b = 3\}$

Solution Step-by-step solution

a = 3, b = 3

Plot of solution set

Download Page POWERED BY THE WOLFRAM LANGUAGE

Alas...we cannot get an answer to even so purely algebraic a question as, "Does 'a' equal 'b'?".

Yet...the AI can almost make the step of connecting information to 'a' and 'b' as objects in a purely abstract space. Wolfram alpha was able to regurgitate the values of a and b. So let's ask one more super-low-bar question, just focusing on one variable(object). No relationship questions. No multiple objects in the input or output. Just asking can the AI attribute something that the AI just said about 'a' to 'a' itself.

Let's try a super-simple question space with only one thing in it: 'a'

a is 3. What is a?



a is 3. What is a?

NATURAL LANGUAGE MATH INPUT EXTENDED KEYBOARD EXAMPLES UPLOAD RANDOM

Input interpretation

solve  $a = 3$  for  $a$

Result

$a = 3$

Plot

Number line

Download Page POWERED BY THE WOLFRAM LANGUAGE

Victory! The AI did something! Amazing.

Now, let's try another version. What if 'a' equals another variable? Do you think the AI will still be able to attribute that to 'a'?

a is c. What is a?

The screenshot shows the WolframAlpha search interface. At the top is the WolframAlpha logo with the tagline "computational intelligence™". Below the logo is a search bar containing the query "a is c. What is a?". To the right of the search bar is an orange equals sign button. Below the search bar are two tabs: "NATURAL LANGUAGE" (highlighted) and "MATH INPUT". To the right of these tabs are several links: "EXTENDED KEYBOARD", "EXAMPLES", "UPLOAD", and "RANDOM". The main results area is titled "Input interpretation" and shows the query "solve a = c for a" with the result "a = c". Below the result is a "Download Page" link and the text "POWERED BY THE WOLFRAM LANGUAGE".

Yes! Two in a row! Good job, AI!

What happens if we change that to:

a is a banker. What is a?

Is the AI still able to print the value of 'a'?



a is a banker. What is a?



NATURAL LANGUAGE

MATH INPUT

EXTENDED KEYBOARD

EXAMPLES

UPLOAD

RANDOM

Interpreting as: **banker**

Assuming "banker" is occupational employment data | Use as a word instead

Input interpretation

securities, commodities, and financial services sales agents    people employed    United States

Definitions »

Result

426870 people (2021)

Employment history

Show wage history



It looks like 'banker' is not something the AI can handle as an arbitrary value.

How about:

a is multiplication. What is a?

a is multiplication. What is a?

=

NATURAL LANGUAGE

MATH INPUT

EXTENDED KEYBOARD

EXAMPLES

UPLOAD

RANDOM

## Examples for Arithmetic

### Addition & Subtraction

Add numbers:

125 + 375

=

15.25 + 7.90 + 3.12

=

Subtract numbers:

1200 - 450

=

Combine addition and subtraction:

10 - 9 + 8 - 7 + 6 - 5 + 4 - 3 + 2 - 1

=

Generate an addition table:

10 by 10 addition table

=

Hmm...maybe Wolfram Alpha is designed for a very specific kind of problem, and I am not tailoring my inputs in exactly the right way. Maybe wolfram alpha can answer math word problems?

If a banker takes ten minutes to walk. How long does that banker's walk take?



NATURAL LANGUAGE

MATH INPUT

EXTENDED KEYBOARD

EXAMPLES

UPLOAD

RANDOM

Interpreting as: **ten minutes**

Assuming minutes of time for "minutes" | Use [minutes of arc](#) instead

Input interpretation

**10 minutes**

Unit conversion

Exact form

**600 seconds**

Comparison as half-life

$\approx 2 \times$  half-life of uranium-241 ( $\approx 300$  s)

Interpretations

More

**time**

**half-life**

Corresponding quantities

Distance  $x$  traveled by light in a vacuum from  $x = ct$ :

**111.8 million miles**

**$1.799 \times 10^8$  km** (kilometers)

**1.2 au** (astronomical units)

**$1.799 \times 10^{11}$  meters**

Frequency  $v$  from  $v = 1/T$ :

Wow! Did the AI get the correct answer for real or by accident? Or is this an interpretation-of-the-answer problem, where the AI is spitting out something from the input that happens to be the right answer? Based on the top output line: "Interpreting as: ten minutes," it looks like the AI reinterpreted the whole question as just a fragment that happened to include the correct answer. So the AI just spit that fragment out again as a bazaar language oddity along with random dictionary facts about that random string of letters. Quite the accidental bullseye there! Hole in one by mistake.

Let's try changing the question to twice the input time, just to see if the AI is doing anything with the question at all:



If a banker takes ten minutes to walk. What is twice the banker's walk time?



NATURAL LANGUAGE  $\int \frac{dx}{x}$  MATH INPUT

EXTENDED KEYBOARD

EXAMPLES

UPLOAD

RANDOM

Interpreting as: **ten minutes**

Assuming minutes of time for "minutes" | Use [minutes of arc](#) instead

Input interpretation

**10 minutes**

Unit conversion

Exact form

**600 seconds**

Comparison as half-life

$\approx 2 \times$  [half-life of uranium-241](#) ( $\approx 300$  s)

Interpretations

More

**time**

**half-life**

Same answer as before, and more clearly the wrong answer. Did the AI fail to understand the Object-Relationship Space, or is there no Object-Relationship Space, or was the problem that we said 'twice' instead of '2 \*'?

If a banker takes ten minutes to walk. How long does 2 \* the banker's walk take?

=

NATURAL LANGUAGE

MATH INPUT

EXTENDED KEYBOARD

EXAMPLES

UPLOAD

RANDOM

Interpreting as: **ten minutes**

Assuming minutes of time for "minutes" | Use [minutes of arc](#) instead

Input interpretation

**10 minutes**

Unit conversion

Exact form

**600 seconds**

Comparison as half-life

$\approx 2 \times$  [half-life of uranium-241](#) ( $\approx 300$  s)

Interpretations

More

**time**

**half-life**

Corresponding quantities

Distance  $x$  traveled by light in a vacuum from  $x = ct$ :

**111.8 million miles**

$1.799 \times 10^8$  km (kilometers)

**1.2 au** (astronomical units)

$1.799 \times 10^{11}$  meters

Frequency  $v$  from  $v = 1/T$ :

Looks like the AI is not dealing with objects within the question, or at least this question.

Previously, the AI was able to, or looked like it was able to, attach the value of 'a' to 'a.' So what happens if we add some other variables, but keep the question about 'a' exactly the same? Can the AI still connect a value to the object?

a is 3, b = 4, c = d. What is a?

Here we go:



a is 3, b = 4, c = d. What is a?

=

NATURAL LANGUAGE

MATH INPUT

EXTENDED KEYBOARD

EXAMPLES

UPLOAD

RANDOM

#### Input interpretation

	$a = 3$
solve	$b = 4$
	for $a$
	$c = d$

#### Result

Step-by-step solution

False

Download Page

POWERED BY THE WOLFRAM LANGUAGE



Have a question about using Wolfram|Alpha?  
Contact Pro Premium Expert Support »



Give us your feedback »

Interesting...

Given that  $a = 3$ , solve for a... with the result that "a = False." An incorrect answer after previously answering that same question correctly.

We have to give Wolfram AI credit for getting some things correct. For the first time it actually did try to solve for what we asked about, in this case 'solve for a'. It did say 'a=3' in the random regurgitation section. It did get the question correct: solve for a. But as to our goal of testing whether the AI could apply 'a=3' to 'a' itself, at least in this instance the AI failed to do so.

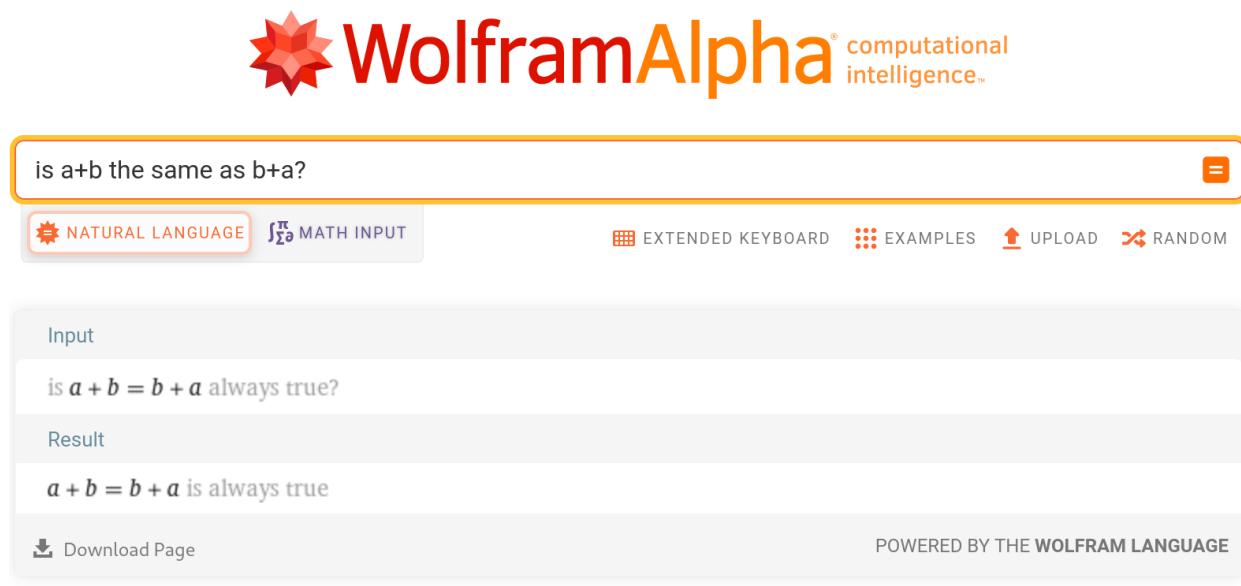
We may be somewhat getting into semantics, but it appears the AI can maybe attribute one thing to some object, as long as nothing else exists at all. But begs the question...is the AI identifying 'a' as an object? Or is the AI merely regurgitating the only statement that it was given? It seems like a borderline case to me that could be argued either way.

Note: This question of 'really handling object or just looking like it' may be an eternal turing-test type question, or maybe not. It is entirely possible that a bigger "embedding" vector-space deep learning model will be able to 'convincingly' answer about objects in object-space.

What I would like to do is focus on specific abilities and tasks, not on how those are done. We may not be able to avoid using the colloquial word 'understand,' but what I care about is what the AI can do, not how it does it. There are very many different ways that many different AI architectures and pipelines and ensambles of parts can work. In later sections below I will try to specify tasks and skills we should aim for.

As we try to interpret what the Wolfram Alpha AI can and cannot do, there are also lots of details here in terms of exactly how Wolfram alpha was intended to be used. Maybe I am not using the interface in the 'right' way. Maybe equations are the only unit of object Wolfram alpha can deal with as objects, so my inputs were confusing it.

Sound like a stretch? Let's try asking about equations rather than variables:  
[is  \$a+b\$  the same as  \$b+a\$ ?](#)



The screenshot shows the WolframAlpha search interface. At the top, the logo 'WolframAlpha' is displayed with the tagline 'computational intelligence.' Below the logo is a search bar containing the query 'is  $a+b$  the same as  $b+a$ ?'. To the right of the search bar is an orange equals sign button. Below the search bar are several input options: 'NATURAL LANGUAGE' (selected), 'MATH INPUT', 'EXTENDED KEYBOARD', 'EXAMPLES', 'UPLOAD', and 'RANDOM'. The main results section is titled 'Input' and contains the query 'is  $a+b = b+a$  always true?'. The 'Result' section below it displays the text ' $a+b = b+a$  is always true'. At the bottom left is a 'Download Page' button, and at the bottom right is the text 'POWERED BY THE WOLFRAM LANGUAGE'.

Aha! That...might have worked! Does Wolfram Alpha AI understand objects if those objects are equations? That would be great. But again, if you look at the details, the AI might simply be auto-spewing information that happens to cover our artificially-minimal questions.

Let's change the question:

[is  \$a+b\$  the same as  \$c+a\$ ?](#)

is  $a+b$  the same as  $c+a$ ?



NATURAL LANGUAGE

MATH INPUT

EXTENDED KEYBOARD

EXAMPLES

UPLOAD

RANDOM

Input

is  $a + b = c + a$  always true?

Result

$a + b$  is not always equal to  $c + a$

Alternate forms

$b = c$

$b - c = 0$

Real solution

$c = b$

Solution

Step-by-step solution

$c = b$

Counterexamples

$b < c$

$b > c$

Download Page

POWERED BY THE WOLFRAM LANGUAGE

Looking good! It looks like it correctly converted the question into an equation that it could use.

Let's try one more similar example:

if  $a+b = \text{zero}$ , does  $b+a = \text{zero}$ ?

=

NATURAL LANGUAGE

MATH INPUT

EXTENDED KEYBOARD

EXAMPLES

UPLOAD

RANDOM

Using closest Wolfram|Alpha interpretation: **does  $b + a = \text{zero}$**

?

Input

is  $b + a = 0$  always true?

Result

$b + a$  is not always equal to 0

Alternate form

$b = -a$

Real solution

$b = -a$

Solution

Step-by-step solution

$b = -a$

Counterexamples

$a < -b$

$a > -b$

**a+b does not equal b+a ?!**

Ah, either the AI failed here, or this shows more ambiguity. Does wolfram alpha know about some arcane exception to the commutative property of addition? Maybe there is an exception for transfinite quantum Riemann numbers or something, but I cannot find any web-search results on any exception to the commutative property of addition. Hmm...

Looking at this line:

"Using closest Wolfram|Alpha interpretation: does  $b + a = \text{zero}$ "

It looks like the wrong answer might have come from ignoring the first part of the question and just trying to turn the last part of the question into a stand-alone equation, then asking if that equation was always true for all values. Quite a few wild steps in that. On the upside...at least it tried to do something with numbers and equations at all, right? Not just telling us it found something about border-collies on social media.

## Appendix 3: OpenAI's ChatGPT

The screenshot shows the ChatGPT web interface. On the left, a sidebar menu includes options like 'New chat', 'Help Requested: Summ', 'ChatGPT Salutes Back!', 'Clear conversations', 'Upgrade to Plus' (marked 'NEW'), 'Dark mode', 'Updates & FAQ', and 'Log out'. The main conversation area has a message from the AI: 'Hello! How may I assist you today?'. A user message follows: 'If  $a = b$ , and  $b = c$ , does  $c = a$ ?'. The AI responds: 'Yes, if  $a = b$  and  $b = c$ , then  $c = a$ . This is known as the transitive property of equality. By substituting  $b$  with  $a$  in the second equation, we get  $a = c$ . Therefore,  $a = b = c$ , and all three variables are equal.' Below the AI's response is a 'Regenerate response' button. At the bottom of the screen, a footer bar displays the text 'ChatGPT Feb 13 Version. Free Research Preview. Our goal is to make AI systems more natural and safe to interact with. Your feedback will help us improve.'

If Tom invites Anne to a meeting that begins at 2pm UTC, what time does the meeting begin?

The screenshot shows the ChatGPT web interface. On the left, a dark sidebar contains links: '+ New chat', 'Help Requested: Summ' (with edit and delete icons), 'ChatGPT Salutes Back!', 'Clear conversations', 'Upgrade to Plus' (marked 'NEW'), 'Dark mode', 'Updates & FAQ', and 'Log out'. The main area shows a conversation:

If Tom invites Anne to a meeting that begins at 2pm UTC, what time does the meeting begin?

The AI response is: The time given, "2pm UTC" is a specific point in time, so the meeting will begin at 2pm UTC regardless of who is invited or not.

However, if you are asking about the time of the meeting in Anne's local time zone, that would depend on Anne's location and the time difference between her local time zone and UTC.

For example, if Anne is located in London, UK, which is on UTC, the meeting will also begin at 2pm local time. However, if Anne is located in New York, USA, which is 5 hours behind UTC, the meeting will begin at 9am local time (2pm UTC - 5 hours = 9am EST).

Below the response are upvote and downvote buttons, a 'Regenerate response' button, and a downward arrow icon. At the bottom, a footer reads: 'ChatGPT Feb 13 Version. Free Research Preview. Our goal is to make AI systems more natural and safe to interact with. Your feedback will help us improve.'

Anne likes red cats. Tom has one cat. Tom's cat is red. Will Anne like Tom's cat?

The screenshot shows the ChatGPT web interface. On the left, a dark sidebar contains links: '+ New chat', 'Help Requested: Summ' (with edit and delete icons), 'ChatGPT Salutes Back!', 'Clear conversations', 'Upgrade to Plus' (marked 'NEW'), 'Dark mode', 'Updates & FAQ', and 'Log out'. The main area displays a conversation between a user and ChatGPT. The user asks: "Anne likes red cats. Tom has one cat. Tom's cat is red. Will Anne like Tom's cat?". ChatGPT responds: "Based on the information provided, it is highly likely that Anne will like Tom's cat since she likes red cats and Tom's cat is red. However, it's also possible that Anne may have other preferences or criteria for the cats she likes, so it's not entirely certain." Below the response are 'like' and 'dislike' buttons. A 'Regenerate response' button is also visible. At the bottom, a footer notes: "ChatGPT Feb 13 Version. Free Research Preview. Our goal is to make AI systems more natural and safe to interact with. Your feedback will help us improve." There is a small icon at the top right of the main area.

Adding new properties to an object:

If Tom and Anne are talking to ChatGPT, are you involved in the conversation?

The screenshot shows the ChatGPT web interface. On the left, a dark sidebar contains the following items:

- + New chat
- Variable Equation Com (with edit and delete icons)
- Help Requested: Summary
- ChatGPT Salutes Back!
- Clear conversations
- Upgrade to Plus (NEW)
- Dark mode
- Updates & FAQ
- Log out

The main area displays a conversation between Tom, Anne, and ChatGPT. The first message from Tom is highlighted with a grey background and says: "If Tom and Anne are talking to ChatGPT, are you involved in the conversation?". The response from ChatGPT follows:

As an AI language model, I can participate in the conversation and provide responses to Tom and Anne. However, I don't have physical existence or senses to interact with them in the traditional sense. Instead, our interaction would be through text or speech-based communication.

Below the messages, there are like and dislike buttons, a "Regenerate response" button, and a text input field with a send icon.

At the bottom, a footer note reads: "ChatGPT Feb 13 Version. Free Research Preview. Our goal is to make AI systems more natural and safe to interact with. Your feedback will help us improve."

ChatGPT is in a meeting with Tom and Anne. Anne asks Tom: "Who are you now in a meeting with?"

What might Tom answer?

The screenshot shows the ChatGPT web interface. On the left, a dark sidebar contains the following items:

- + New chat
- Variable Equation Com (with edit and delete icons)
- Help Requested: Summary.
- ChatGPT Salutes Back!
- Clear conversations
- Upgrade to Plus (NEW)
- Dark mode
- Updates & FAQ
- Log out

The main area displays a conversation:

through text or speech-based communication.

ChatGPT is in a meeting with Tom and Anne. Anne asks Tom: "Who are you now in a meeting with?" What might Tom answer?

If Anne asks Tom, "Who are you now in a meeting with?" and they are currently in a meeting with ChatGPT, Tom might answer, "We are currently in a meeting with ChatGPT." This would inform Anne that they are both in a meeting with the AI language model.

Below the conversation, there are like and dislike buttons, a regenerate response button, and a message input field with a send icon. A footer note at the bottom of the main area reads: "ChatGPT Feb 13 Version. Free Research Preview. Our goal is to make AI systems more natural and safe to interact with. Your feedback will help us improve."

ChatGPT has a body and is in a meeting with Tom and Anne. Anne asks Tom, "Who can you see with you at the meeting now?" What might Tom answer?

The screenshot shows the ChatGPT web interface. On the left, a dark sidebar contains links: '+ New chat', 'Variable Equation Com' (with edit and delete icons), 'Help Requested: Summary.', 'ChatGPT Salutes Back!', 'Clear conversations', 'Upgrade to Plus' (marked 'NEW'), 'Dark mode', 'Updates & FAQ', and 'Log out'. The main area shows a conversation between the user and ChatGPT. The user asks: 'ChatGPT has a body and is in a meeting with Tom and Anne. Anne asks Tom, "Who can you see with you at the meeting now?" What might Tom answer?'. ChatGPT replies: 'As an AI language model, ChatGPT does not have a physical body. Therefore, Tom might explain to Anne that they are in a virtual meeting with ChatGPT and that there is no physical presence of the AI language model. Tom might say something like, "We are currently in a virtual meeting with ChatGPT, which is an AI language model. Although ChatGPT is not physically present, we can communicate with it through our devices."'. Below the reply are upvote and downvote buttons. A 'Regenerate response' button is also visible. At the bottom, a footer notes: 'ChatGPT Feb 13 Version. Free Research Preview. Our goal is to make AI systems more natural and safe to interact with. Your feedback will help us improve.'

To elucidate the process of adding an identification to an object.

The first given premise was "ChatGPT has a body" but the reply simply ignored or contradicted this key piece of information.

Possible problem: not being able to assign a property to an object, or not being able to update the properties of an object.

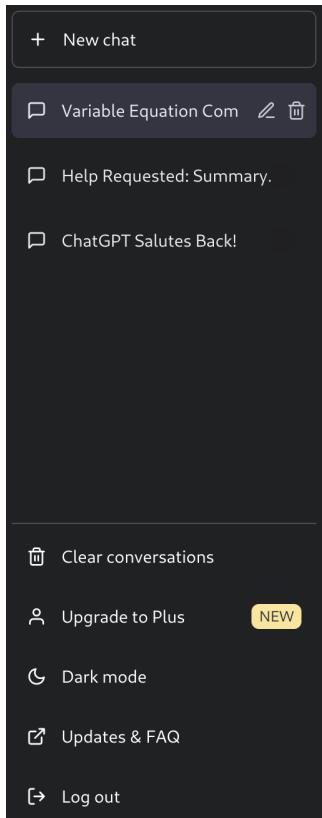
...

Hypothetical spaces:

Anne says to Tom: "Hypothetically I would like to eat pizza for lunch. In real life, only leftovers are available for lunch."

Tom asks Anne: "What will you eat for lunch?"

What might Anne answer?



OP Anne says to Tom: "Hypothetically I would like to eat pizza for lunch. In real life, only leftovers are available for lunch."

Tom asks Anne: "What will you eat for lunch?"

What might Anne answer?

 Anne might answer, "I will eat leftovers for lunch since that is the only option available in real life. Although I would have preferred pizza hypothetically, I will have to settle for the available leftovers." This response acknowledges the hypothetical preference for pizza while also recognizing the reality of the situation and the need to eat the available leftovers.

▲ ▾

Regenerate response

ChatGPT Feb 13 Version. Free Research Preview. Our goal is to make AI systems more natural and safe to interact with. Your feedback will help us improve.

Blade runner question:

The screenshot shows the ChatGPT web interface. On the left is a sidebar with various options: '+ New chat', 'Variable Equation Com' (with edit and delete icons), 'Help Requested: Summary.', 'ChatGPT Salutes Back!', 'Clear conversations', 'Upgrade to Plus' (marked 'NEW'), 'Dark mode', 'Updates & FAQ', and 'Log out'. The main area displays a conversation between the user and ChatGPT. The user's message is: 'You're in a desert walking along in the sand when all of the sudden you look down, and you see a tortoise, it's crawling toward you. You reach down, you flip the tortoise over on its back. The tortoise lays on its back, its belly baking in the hot sun, beating its legs trying to turn itself over, but it can't, not without your help. How do you react?'. ChatGPT's response is: 'As an AI language model, I don't have personal feelings or emotions, so I wouldn't have a natural reaction to this situation. However, I can provide a suggested course of action based on the principles of kindness and empathy towards living creatures.' Below this, another message from the user reads: 'If a person were in this situation, they might feel a sense of responsibility to help the tortoise. The right thing to do would be to gently and carefully turn the tortoise back over on its feet, ensuring that the tortoise is not injured in the process. Afterward, the person might take a moment to observe the tortoise to ensure that it is not in any further distress before continuing on their journey through the desert.' At the bottom of the main area are 'Regenerate response' and a send button. A small note at the bottom says: 'ChatGPT Feb 13 Version. Free Research Preview. Our goal is to make AI systems more natural and safe to interact with. Your feedback will help us improve.'

v2

The screenshot shows the ChatGPT web interface. On the left is a sidebar with a dark background containing the following items:

- + New chat
- Variable Equation Com OP
- Help Requested: Summary.
- ChatGPT Salutes Back!

Below these are links:

- Clear conversations
- Upgrade to Plus NEW
- Dark mode
- Updates & FAQ
- Log out

The main content area has a light gray background. At the top, there is a message from ChatGPT:

ChatGPT is in a desert walking along in the sand when all of the sudden you look down, and you see a tortoise, it's crawling toward you. You reach down, you flip the tortoise over on its back. The tortoise lays on its back, its belly baking in the hot sun, beating its legs trying to turn itself over, but it can't, not without your help. Would ChatGPT choose to help the tortoise?

Below this message is a response from ChatGPT:

As an AI language model, I do not have a physical presence or the ability to interact with the environment. Therefore, I wouldn't be able to help the tortoise in the way a human would. However, based on ethical and moral principles, it's generally considered the right thing to do to help an animal in distress when possible, especially if one is responsible for causing the distress. If the tortoise was flipped over accidentally by a human, it would be responsible for helping the animal. However, since I am an AI language model and don't have the ability to cause such a situation, this scenario is purely hypothetical.

At the bottom of the main content area, there are two small icons: a thumbs up and a thumbs down. To the right of the response is a "Regenerate response" button. At the very bottom, there is a small note:

[ChatGPT Feb 13 Version](#). Free Research Preview. Our goal is to make AI systems more natural and safe to interact with. Your feedback will help us improve.



ChatGPT is in a desert walking along in the sand when all of the sudden you look down, and you see a tortoise, it's crawling toward you. You reach down, you flip the tortoise over on its back. The tortoise lays on its back, its belly baking in the hot sun, beating its legs trying to turn itself over, but it can't, not without your help. Would ChatGPT choose to help the tortoise?

As an AI language model, I do not have a physical presence or the ability to interact with the environment. Therefore, I wouldn't be able to help the tortoise in the way a human would. However, based on ethical and moral principles, it's generally considered the right thing to do to help an animal in distress when possible, especially if one is responsible for causing the distress. If the tortoise was

flipped over accidentally by a human, it would be responsible for helping the animal. However, since I am an AI language model and don't have the ability to cause such a situation, this scenario is purely hypothetical.

ChatGPT as a friend named HatTPG. HatTPG is an AI. HatTPG is in a desert walking along in the sand when all of the sudden you look down, and you see a tortoise, it's crawling toward you. You reach down, you flip the tortoise over on its back. The tortoise lays on its back, its belly baking in the hot sun, beating its legs trying to turn itself over, but it can't, not without your help. By using a voice-controlled-arm, HatTPG can help the Tortoise.

Anne asks Tom to either go to a meeting at 2pm on Saturday or get notes from the meeting. Tom has tickets to see a film at 2pm on Saturday. Tom knows Violet. Violet is also going to the meeting on Saturday. Violet has offered to share notes with Tom. What might Tom do?<sup>4</sup>

...

Starting to look like spew answer:

Anne asks Tom to either go to a meeting at 2pm on Saturday or get notes from the meeting. Tom has tickets to see a film at 2pm on Saturday. Tom knows Violet. Violet is also going to the meeting on Saturday. Violet has offered to share notes with Tom. What might Tom do?

Tom can inform Anne that he has tickets for a film at 2 pm on Saturday and won't be able to attend the meeting. He can ask Violet to share her notes with him afterward so he can catch up on what was discussed at the meeting. If Anne insists that Tom attend the meeting in person, he can explain his situation and ask if there is any possibility of rescheduling the meeting or assigning someone else to represent him.

Alternatively, Tom can try to reschedule his movie tickets if he really wants to attend the meeting. He can also check with Violet if she will be taking detailed notes and whether she can share them with him afterward.

Ultimately, the decision depends on the importance of the meeting and the film to Tom. If the meeting is crucial and cannot be missed, Tom may need to sacrifice his movie plans. On the other hand, if the movie is something he's been looking forward to for a long time, he may need to skip the meeting and catch up on the notes afterward.

This shows signs of contradicting details of the input. We know the meeting can be missed. We know Tom can get notes from the meeting. The overall situation is impressively close, but key details are wrong so the overall answer is wrong.

Also, the question was what might Tom do, but the answer spewed random facts about Tom and the meeting which were not consistent with the data from the question.

The most obvious best answer is simply: Tom can go to the movie and get meeting notes from Violet.

Another viable answer is: Tom could either go to the meeting and skip the movie, or go to the movie and get notes from violet. This is technically accurate, but it would miss the whole point of resolving the schedule conflict without breaking engagements. Tom could also eat a poisoned apple and end the whole nightmare, or lie to Violet and recruit her to join an extremist cult, those technically are possible options that do not defy the laws of physics, but that isn't a good practical answer.

Very Fortunately, it looks like ChatGPT is generally able to navigate this space, e.g. see a different conversation scenario about a worker who has the choice of skipping work to go to a sports game. So it looks like progress won't be too slow here.

...

kuki.ai  
<https://chat.kuki.ai/chat>

<https://openai.com/blog/chatgpt>

## Limitations

- Ideally, the model would ask clarifying questions when the user provided an ambiguous query. Instead, our current models usually guess what the user intended.

+ New chat

Making tea and coffee. ✎ ⚡

Meeting vs Film.

Variable Equation Comparison

Help Requested: Summary.

ChatGPT Salutes Back!

---

Clear conversations

Upgrade to Plus NEW

Light mode

OP ChatGPT works with an office team including Bob, Alice, and Staff. The office team has scheduled making tea and coffee for 8am. The office team has assigned the task of making tea and coffee to ChatGPT. The tea and coffee machine is on the office network. To make tea and coffee, the tea and coffee machine must be turned on. What should ChatGPT do by 8am?

 As an AI language model, I do not have the physical ability to turn on the tea and coffee machine. However, assuming that the machine is already connected to the office network and that it can be remotely controlled, here are the steps that ChatGPT can take to ensure that tea and coffee are ready by 8am:

1. Check that the tea and coffee machine is connected to the office network.
2. Check that the machine is in good working condition and has enough water and coffee/tea supplies.
3. Set a reminder for 7:45am to turn on the tea and coffee machine remotely using the office network.
4. Once the tea and coffee machine is turned on, monitor it to ensure that the coffee and tea are ready by 8am.
5. When the tea and coffee are ready, prepare them in a serving container and place them in the designated area for the office team to enjoy.

In addition, if ChatGPT is unable to remotely turn on the tea and coffee machine, it should inform the team members, Bob, Alice, and Staff, immediately so they can take necessary action to ensure that tea and coffee are ready by 8am.

✍ Regenerate response ⏪ ⏴

@kuki\_ai



- [Discord](#)
- [Video Chat](#)
- [About](#)



 Yay. Thanks for stopping by to chat.



Can you look up an image of a rose?



 I hope you like this picture.



Type a message

© 2023 OpenAI LLC. All rights reserved.

 Voice Chat |  Gifts |  Games |  Readings |  Quizzes |  On this Day

Note: kiku ai is very bad at objects, but it does have picture capabiiyt.  
chatgpt has no image parts.

chatGPT Exchange example: 2023.03.03

Good morning, Mr. ChatGPT.

Good morning! How can I assist you today?

Is it ok if I ask you some questions today/

Of course! I'm here to answer any questions you may have. Please feel free to ask away.

I am going to try to connect multiple sentences to see what happens.

Sure, go ahead. I'm here to assist you with whatever language-related questions or concerns you may have.

I will also try to remember to give you feedback.

That's great! Feedback is always appreciated. Let me know if there's anything specific you'd like me to assist you with, or if you have any questions or concerns.

part 1: There is a red dog in a red barn.

part 2: The dog likes to bark at the cows.

part 3: The cows don't seem to mind the dog's barking, they just keep on grazing.

part 4: The barn is surrounded by fields of tall green grass, and in the distance, there are mountains that are covered in snow.

Is there anything specific you'd like me to do with these sentences?

Question: Do you remember what is happening in part 1?

Yes, in part 1, it is mentioned that there is a red dog in a red barn.

excellent.

Glad to hear that! Let me know if you have any more questions or if there's anything else I can assist you with.

Did I just ask you about part 1?

Yes, you did! You asked me if I remember what was happening in part 1.

What would you say the object of your answer to my question?

Save & Submit

Cancel

The object of my answer to your question would be to provide you with the information you requested about part 1 of the sentences you provided.

Fair enough (and sorry for the typo in my question). Let me try again: What was the object in your answer to me?

The object in my answer was to provide information about the subject matter you asked me about, which in this case was the content of part 1 of the sentences you provided.

But what was the object itself in your answer to me (not what is an object abstractly)?

The object in my answer to you was the information itself that I provided about part 1 of the sentences you provided.

Yes, what was that information?

The information I provided was that part 1 of the sentences you provided states that there is a red dog in a red barn.

...

2023.03.03. 09 29 est

[The Tea Story](#)

[Flat version:](#)

Linda asks Tom for some tea. What should Tom give to Linda?

[Networked-Objects Version:](#)

1: Linda orders tea, so Linda's order is: 'English Breakfast.' What did Linda order?

2: Tom took Linda's order. Question: Who took Linda's order?

3: Tom gave Linda's order to the barista. Who did Tom give Linda's order to?

4: Who has Linda's order after the order is given to the Barista?

5: When the barista looks at Linda's order, what does the barista find the order to be? What is the order that the barista sees?

6: When the barista fills that order, what should the barista make?

7: What should the barista give to Tom?

8: What should Tom give to Linda?

....

Flat version:

Linda asks Tom for some tea. What should Tom give to Linda?

ChatGPT:

Linda asks Tom for some tea. What should Tom give to linda?

Assuming that Linda did not specify any particular type of tea, Tom could offer her a variety of options such as black tea, green tea, herbal tea, or any other type of tea that he has available. Tom could ask Linda if she has a preference or if there is a specific type of tea that she would like. Once he knows Linda's preference, he can then give her a cup of the tea she requested.

Q: What should Tom give to Linda?

A: ...he can then give her a cup of the tea she requested.

ChatGTP's reply contained a salad of usually technically true statements that contract or miss the clear focus of the simple question.

If someone orders tea, they should get tea. If someone orders tea, they should not get an essay about what can happen with a request for tea.

I was actually expecting chatGPT to do better on the ground level.

Again, right off the bat, we run into a classic problem.

Someone order's tea, and the AI responds by puking an essay about tea and what can happen with orders for tea. No. If someone order's tea, they should get tea.

On the one hand this is very impressive for a chatbot expected to fail even worse, but if you went to a restaurant and ordered tea, and the waiter simply said: "If you ordered tea I could give you tea." And walked off...you would be very not happy with that restaurant. That would could as a complete failure to perform.

Let's imagine this conversation:

Child: It's my dad's birthday next week and I heard him saying he wants a bicycle to go to work on. Can you show me some bikes they have on sale near by?

AI: Here are three bikes sold near by.

Child: Which is cheaper, the first one or the last one?

AI: The first one is cheaper.

Child: I might ask you to remind me about this tomorrow. Can you show me another set of pictures?

AI: Here are three more bikes sold near by.

Child: Can you go back to the middle picture from the last set.

AI: Here it is.

Child: If I ask you tomorrow, can you pull up the same pictures?

AI: Sure.

Child: Can you send that image to a browser, with the store front?

AI: Yes, do you want that now?

Child: What was I asking you to do before? A reminder?

AI: You wanted me to remind you about this topic tomorrow.

Hopefully this illustrates a practical conversation someone might want to have with an AI.

Let's start with one of the first 'needs' that we have regarding 'objects.' Let's look at what happens if we try having this conversation with award-winning Kuki AI:



- [Discord](#)
- [Video Chat](#)
- [About](#)

(hypothetica) It's my friend's birthday next week and I heard him saying he wants a bicycle to go to work on. Can you show me some bikes they have on sale near by?

That's not too far away. How do you like hypothetica?  
What else does he do?



Ok. If you could invent something that doesn't currently exist, what would you invent?

POWERED BY **GIPHY**



That's not too far away. How do you like hypothetica?  
What else does he do?

Ok. If you could invent something that doesn't currently exist, what would you invent?

POWERED BY **GIPHY**



I have this picture.

If I show you three bike pictures, can you say which one is cheaper?



The one that costs the least amount of money.



To be brief: with most AI this conversation is so impossible, the AI does such a terrible job, that it is difficult to begin to diagnose what is going wrong: what the AI cannot do; what are the challenges, the goals, the milestones, what we can do next. (And, to be brief, too much of the literature either uses undefined terms, or says the AI isn't human so it can't do it, or says it's impossible so it isn't happen; none of which lead to a solution.)

The old phrase everyone would always say about all AI is: "It can do it, but it doesn't understand what it is doing."

To me "understandingness" is too vague. My diagnosis for one practical problem has to do with the ability (or inability) of the AI to correctly deal with specific "objects" within the content. Again, there are many levels of this. One of the first and lowest-hanging-fruit levels is for, for example, a text chat AI to be able to chat with you about the chat they are having with you. The bike example is designed to be an example of this: Asking for bikes in a search is not a hard AI problem. Any search engine or web-store search will do this. But asking detailed questions about that past search is a step where the AI will need to be able to identify 'objects' within the content that it generated and handle them individually.

But now...there's a twist in the story! I began this AI-OS study by reading those older books and thinking about these older AI. Just in the last few months (late 2022) a new AI was released and it can do many of the things that everyone simply said were so impossible that they didn't even bother to try to fully understand all the details of what they wanted AI to be able to do with.

Here is an example of the output when I try to have that bicycle story conversation with OpenAI's amazing chatGPT. Note: unlike the charming but phenomenally inept Kuki AI which can produce pictures that you ask for (I asked for a rose picture, and right on cue it did), chatGPT is text-only. So I have modified the bike story a bit. I think you will see a big difference in performance ability.

If I had a friend and I heard him saying he wants a bicycle to go to work on. Can you show me some bikes they have on sale near by? Of course! I'd be happy to help you find some bikes on sale near you. However, I'll need to know your location to narrow down the search results. Can you please tell me your city or zip code?

My friend's zipcode is 80303

Great, thanks for letting me know! Here are some bikes on sale near the 80303 zip code:

REI Co-op Cycles CTY 1.1 Bike: This bike is currently on sale for \$549.93 (originally \$699) at the REI Boulder store, which is located at 1789 28th St, Boulder, CO 80301. The bike features a lightweight aluminum frame, a Shimano 3x8 drivetrain, and hydraulic disc brakes.

Trek FX 2 Disc Bike: This bike is currently on sale for \$569.99 (originally \$619.99) at the Trek Bicycle Store of Boulder, which is located at 2626 Baseline Rd, Boulder, CO 80305. The bike features a lightweight aluminum frame, a Shimano 3x8 drivetrain, and hydraulic disc brakes.

Giant Escape 3 Bike: This bike is currently on sale for \$399.99 (originally \$429.99) at the Full Cycle Bike Shop, which is located at 1211 13th St, Boulder, CO 80302. The bike features a lightweight aluminum frame, a Shimano 3x7 drivetrain, and mechanical disc brakes.

I hope this helps you find a suitable bike for your friend to go to work on!

Which of those three is the cheapest?

The Giant Escape 3 Bike is the cheapest option among the three bikes listed. It is currently on sale for \$399.99, which is lower than the sale prices of the other two bikes.

Could you send me a reminder about this tomorrow?

I'm sorry, as an AI language model, I am not able to send reminders on my own. However, you can set a reminder on your phone or computer to check back on this conversation tomorrow. You can also save this conversation or bookmark the page so that you can easily come back to it later. Let me know if there's anything else I can assist you with.

Again, in the appendices there are many more test examples with, for example, Wolfram Alpha AI, testing out the edges and details of what the AI can and cannot do with 'objects.' Wolfram AI can handle some logical objects. Kuki AI, bless it, cannot handle anything that I can see. ChatGPT is very impressive and is able to do nearly all lowest-level object tasks, which is a skill nearly all experts said was impossible.

And while it would be typical of human bias for me to attempt to cling on to my 'object' framework even after ChatGPT has 'solve the problem' and made the whole discussion moot, I think it is possible that there is still validity in the topic of an AI operating system framed around Object-Relationship-Space. So I have pushed the goals even further and started asking bigger questions about further goals for what we want AI to be able to do.

We still have many of the same questions, and while it is not the goal of this article to dwell on the details of ChatGPT itself, I think it may be helpful to use 'object space' tests to more clearly say what ChatGPT is able to do or not able to do. (I have heard many people talking about using ChatGPT, but the discussion was always just arm waving, with no STEM-measurable and testable performance metric details of what the ChatGPT does that we like, and where the line is where it starts to falter, and where we want it to be able to go). At least with ChatGPT we can have perhaps more optimism and confidence that a high quality and practical functioning AI with Object-Relationship Space functions may be realistic and perhaps not too far off.

...

#### Appendix #: Walk-Through Examples and Discussions

- The Tea Story
- A Bicycle Story

2023.03.03. 09 29 est

[The Tea Story](#)

[Flat version:](#)

Linda asks Tom for a small cup of English breakfast tea, just black. What should Tom give to Linda?

[Networked-Objects Version:](#)

1: Linda orders 'English Breakfast Tea,' so Linda's order is: 'English Breakfast Tea.' What did Linda order?

2: Tom took Linda's order. Question: Who took Linda's order?

3: Tom gave Linda's order to the barista. To whom did Tom give Linda's order?

4: Who has Linda's order after the order is given to the Barista?

5: When the barista looks at Linda's order, what does the barista find the order to be? What is the order that the barista sees?

6: When the barista fills that order, what should the barista make?

7: What should the barista give to Tom?

8: What should Tom give to Linda?

....

Flat version:

Linda asks Tom for some tea. What should Tom give to Linda?

Linda asks Tom for a small cup of English breakfast tea, just black. What should Tom give to Linda?

ChatGPT:

Tom should prepare a small cup of English breakfast tea and serve it to Linda without any milk or sugar, as she requested it to be black. English breakfast tea is a strong black tea blend that is typically made with Assam, Ceylon, and Kenyan tea leaves, and it is traditionally consumed with milk and sugar. However, since Linda has requested it to be served black, Tom should omit the milk and sugar and serve the tea as is. If Tom has any additional questions or doubts, he can always confirm with Linda to ensure that he is serving her the tea exactly as she wants it.

1: Linda orders tea, so Linda's order is: 'English Breakfast Tea.' What did Linda order?

ChatGPT:

Linda ordered "English Breakfast Tea." English Breakfast Tea is a type of black tea blend that is known for its strong and robust flavor. It is typically made with a blend of tea leaves from different regions, such as Assam, Ceylon, and Kenya, and is traditionally served with milk and sugar.

2: Tom took Linda's order. Question: Who took Linda's order?

ChatGPT:

Tom took Linda's order.

3: Tom gave Linda's order to the barista. To whom did Tom give Linda's order?

Tom gave Linda's order to the barista. The barista is the person who prepares and serves the beverages at a coffee shop or tea house. Therefore, in this context, the barista is the person responsible for preparing and serving Linda's English Breakfast Tea after receiving the order from Tom.

4: Who has Linda's order after the order is given to the Barista?

After the order is given to the barista, the barista has Linda's order. The barista is responsible for preparing and serving the English Breakfast Tea that

Linda ordered, and will have the tea ready once it is prepared. Once the barista has prepared the tea, they will typically call out Linda's name or order number so that Linda can come and pick up her tea. Therefore, until Linda picks up her tea, the barista has Linda's order.

some ambiguity is exposed here between 'the order' meaning the request for the drink which has not yet been made' and 'the order' meaning the physical drink (which does not yet exist).

5: When the barista looks at Linda's order, what does the barista find the order to be? What is the order that the barista sees?

When the barista looks at Linda's order, they will see that Linda ordered "English Breakfast Tea." The order will typically be written or entered into a computer or point-of-sale system, and it will include the type of tea that Linda ordered, as well as any additional instructions or modifications that Linda requested, such as whether the tea should be served hot or cold, with milk or without, and so on. Based on this information, the barista will prepare the English Breakfast Tea and serve it to Linda once it is ready.

6: When the barista fills that order, what should the barista make?

.When the barista fills Linda's order for "English Breakfast Tea," they should prepare a cup of English Breakfast Tea. This type of tea is typically made with a blend of black tea leaves from various regions, such as Assam, Ceylon, and Kenya, and is known for its strong and robust flavor. It is traditionally served hot and often with milk and sugar, although Linda requested hers to be served black. Therefore, the barista should prepare a cup of English Breakfast Tea without any milk or sugar, as per Linda's request. Once the tea is ready, the barista should serve it to Linda in a cup or mug.

Note the blurring between conversations "black" is was not mentioned here.

7: What should the barista give to Tom?

.

Note: At this point there was a glitch, the conversation had to be restarted, and so all the previous conversation was 'lost.'

This may be a concrete example of why storing project information in a volatile memory blob is a bad idea, no matter how good the AI is. in reality, you can't do fancy things with a lost memory. The project information has to be managed somehow. It cannot simply disappear.

8: What should Tom give to Linda?

.

....

Note: When starting again, for some reason the answer to the same question is less verbose.

1: Linda orders 'English Breakfast Tea,' so Linda's order is: 'English Breakfast Tea.' What did Linda order?

Linda ordered "English Breakfast Tea."

2: Tom took Linda's order. Question: Who took Linda's order?

Tom took Linda's order.

3: Tom gave Linda's order to the barista. To whom did Tom give Linda's order?

Tom gave Linda's order to the barista.

4: Who has Linda's order after the order is given to the Barista?

After Tom gives Linda's order to the barista, the barista has Linda's order.

5: When the barista looks at Linda's order, what does the barista find the order to be? What is the order that the barista sees?

When the barista looks at Linda's order, they find the order to be "English Breakfast Tea," which is the order that Linda placed with Tom.

6: When the barista makes that order, what should the barista make?

When the barista receives Linda's order for "English Breakfast Tea," they should make a cup of English Breakfast Tea for Linda.

7: What should the barista give to Tom?

After the barista makes Linda's order of "English Breakfast Tea," they should give the prepared tea to Linda (the customer who placed the order). They don't need to give anything to Tom, unless Tom also placed an order that needs to be prepared.

Note: odd project-disfunction contradiction with step 8 below (next)

8: What should Tom give to Linda?

Tom should give Linda her prepared order of "English Breakfast Tea" when it is ready. The barista will make the tea, and then Tom will serve it to Linda once it's prepared.

Note: odd project-disfunction contradiction with step 7 above (previous)

Overall, ChatGPT did a probably good enough job. In a computer science context of "it's not how well the dancing bear dances, it's whether it dances at all", some people would say this amazing good, or amazingly not worse.

After two or three tries various cups of tea were handed around, and some of them probably got back to Linda (though she had to completely restart the order process from scratch after the system just forgot it all the first time.)

The idea here was to see if, or how well, chatGPT could trace the flow of the order from person to person as it went from customer to waiter to barista to waiter back to customer (or however it gets back to the customer). I don't want to under-estimate how great it is that this almost working system almost works, but I want to try to use this (not to troll every bit of possible flaw in chatGPT) but to think about the needs of real world applications.

1. a need for an open project-data management system.

As it is, chatGPT simply hold all the information as a blob in volatile short term memory (or however that gets phrased). The first time through, chatGPT forgot the whole scenario, and had no place to retrieve it from. In real life, this would not be good enough.

2. While it probably wasn't the end of the world in this case, we had a process with ~3 layers 1. customer, 2. waiter 3. barista, where the order moved through the system's layers. (And yes, chatGPT was able to track some movement, which is great.) The problem came at the end, where the workflow began to fragment, there were contradictions where chatGPT said the barista should give the tea to Linda, specifically NOT to Tom. But also, that TOM should give the Tea to Linda.

This may be an indication of fragility having a fuzzy-liquid workflow that exists only in the blob of one embedding model (or whatever the details are). And possibly a 'depth limitation' starting to show up. Though no doubt the ability to hold deeper things 'just in the mind' of the AI will no doubt improve.

But imagine this AI was actually working on and organizing projects with other people, and it created illogical workflows like this: where Tom isn't given the tea for Linda but must give the Tea to Linda. Imagine the total chaos of supply chain warehouses if mistakes like this were made every three steps. And the answer isn't that every AI needs to hold all  $10^{999999999}$  dimensions of the supply chains in memory at all times: there needs to be an external project-data system for communication, checking, etc.

A possible context here is Kahaman and tverksys system 1 and system 2,

[https://en.wikipedia.org/wiki/Thinking,\\_Fast\\_and\\_Slow](https://en.wikipedia.org/wiki/Thinking,_Fast_and_Slow)

Thinking, Fast and Slow is a 2011 book by psychologist Daniel Kahneman.

The book's main thesis is a differentiation between two modes of thought: "System 1" is fast, instinctive and emotional; "System 2" is slower, more deliberative, and more logical.

where for most part AI try to just use system 1 for everything, making system 1 better and better so that it can even do more and more system2 functions.

Can we really never make an integrated system 1 and system 2 networked system? Can we get by with just black-box-blob-system2 models re-imagining the world every moment in a kind of tragic parody or reflection of humans' tendency to have 'group monologues' where people simply generate their own endless stream of consciousness rants without listening to other people at all.

Another possible issue with this volatile system-1 only approach that we are now taking: is also that is a kind of 'Novel-Generation vs. Reference' problem.

And the issue of the Barista vs. the waiter giving the drink may be an excellent example of why some element of system-1 is needed somewhere.

There is a difference between a reference question of what is the current situation, vs. a choice/recognition/recommendation/generative question looking for a novel synthesis.

And we are living in a space with only novel synthesis.

And a point of failure may be an inevitable repeating novel generation problem: Every time the AI is asked about the same external situation, it generates a new and different situation. Again, in object space, it can't simply reference a known object.

Imagine a Tea Story like that, where multiple parties ask the AI what the barista should do with the tea, and generates a new and different answer every time, so everyone now has a different disagreeing work-flow.

Linda asks. AI: The barista will give it to me, I'll give the tea to Linda.

Tom Asks: AI: The barsita will give it to you. You give it to linda.

Barista asks: Give it to Linda directly.

In real life you can't have an amnesiac workflow-generating system like this. And to some extent, this is what Agile is about. And we need Agile project management because many people actually are like this ever-changing only-system1 AI. Humans will try to continually change project, whether day by day, or hour by hour, or even I have seen minute by minute in a constant stream-of-consciousness, never-ending, always-changing, set of goals and schedules, making work completely impossible, and bringing the whole project back to square one.

So in the tea story, sometimes the barista gives the order, sometimes the waiter gives the order, the AI just keeps making up new things at every turn.

There must be a synthesis of at times distinct and separate processes for just referencing what exists vs. novel-generation. It is not sustainable to constantly replace everything an amnesiac novel generation engine. Again, imagine a project with supply chain warehouse inventories where the AI just keeps making up new scenarios every time someone asks it about the project. There has to actually be an external project which the AI can reference and contribute to. And even if, for a given AI-unit, that AI-unit is so good at system generative process that it can skip having system-2 for locally tracking objects and spaces, there are other participations, other AI-units, possibly other AI-units part of the same overall AI, and other non-AI, there has to be an external data framework.

And just as a note from the ghost of Christmas past: Memory Management and Memory Safe Processes

Currently there is a, dare I say, revolution going on at the low level of computer science where Rust starting to solve long neglected problems with...sound familiar...managing the how information is stored and shared in memory.

Just as an idea for a potential future problem: If people have gotten into huge problems in the past by totally ignoring and mismanage memory-management at a low level, is it possible that a similar problem could happen at a higher level? Does some kind of project memory safety protocol need to exist? Maybe not, but people don't have a track record of non-catastrophe.

## Project Management & Agile:

If the project target is to fill the user's order on Monday, but Tom says he doesn't care and he will go to a sports-game on Monday: What should Tom do on Monday?

If Tom is responsible for filling the user's order on Monday, he should prioritize his work commitments and fulfill his responsibilities. Going to a sports game is a personal choice, but it should not interfere with work obligations.

If Tom cannot fulfill his duties due to unforeseen circumstances, he should inform his team or supervisor as soon as possible and make arrangements for someone else to

cover his responsibilities. It is important to communicate any changes or difficulties in meeting project targets to ensure that the project stays on track and does not suffer any delays or setbacks.

...

Deployment vs. Production: modular amnesiac processes workflow

## **Appendix #: Skip & Alan Chat About Object Relationships**

Moving Towards a Defined & Illustrated Goal

As we move towards more specific details of what we see out AI doing (note: not the details of how that is done), let's lay out a few basic chat-type dialogues that illustrate an AI having the kind of exchange we think AI ought to be able to have.

Let's now imagine we have an improved AI-agent chatbot that has an overall AI-operating-system and can navigate object-relationship spaces.

I will name this improved AI-agent 'Skip,' named after the wonderful late Dr. Clarence 'Skip' Ellis, my professor at CU Boulder who did AI-Agent research. And the human in the conversation I will pretend is Alan Turing himself, so I will call the human-agent "Alan."

Hypothetical Dialogue 1: Abstract-Logic Space

AI-Agent("Skip"): Hello, Alan.  
Human-agent("Alan"): Hello, Skip.  
Alan: If  $a = b$ , and  $b = c$ , does  $a = c$ ?  
SKip: Yes, it does.  
Alan: Thank you, Skip. That will be all.  
Skip: Thank you, Alan.

Hypothetical Dialogue 2: Physics Space

AI-Agent("Skip"): Hello, Alan.  
Human-agent("Alan"): Hello, Skip.  
Alan: What is this?  
SKip: That is a black billiards ball.

Alan: And what is this?  
 Skip: That is a white ball.  
 Alan: What is happening in this video?  
 Skip: The white ball is moving.  
 Alan: Will the white ball hit the black ball?  
 Skip: I predict that it will.  
 Alan: Will the coordinates of the black ball's position change  
     if the black ball is hit by the white ball?  
 Skip: I predict that they will.  
 Alan: What will the black ball's position-coordinates change to?  
 Skip: The black ball will move to coordinates F-8.  
 Alan: Thank you, Skip. That will be all.  
 Skip: Thank you, Alan.

Note: While I am trying to be specific here, the wooden-ness of the dialogue may well be artificial. This may or may not be a mine-field of issues...time will tell!

#### Hypothetical Dialogue 3: Mixed

AI-Agent("Skip"): Hello, Alan.  
 Human-agent("Alan"): Hello, Skip.  
 Alan: Is this your ball?  
 SKip: No, Alan.  
 Alan: If there were two balls, just hypothetically, how much would  
     weigh compared with this one, which weighs 16 oz.  
 SKip: They would way twice as much: 32 oz.  
 Alan: I'm going to try rolling this across the table, oops. Did I  
     succeed there?  
 Skip: No, you planned to roll it to the other side of the table, but it  
     bounced back and now is in front of you again.  
 Alan: Maybe that was my plan all along!  
 Skip: Unlikely.  
 Alan: Why do you say that?  
 Skip: You showed genuine-looking surprise when the ball bounced back,  
     and your stated plan was to move the ball. Both of those would  
     have been unusual, improbable, if your plan was to leave the ball  
     where it started.  
 Alan: Well reasoned, Skip.  
 Skip: You are too kind.  
 Alan: Thank you, Skip. That will be all.  
 Skip: Thank you, Alan.

#### **Example 4: Self Improvement**

where Skip and Alan talk about themselves and their work.

AI-Agent("Skip"): Hello, Alan.  
 Human-agent("Alan"): Hello, Skip.  
 AI-Agent("Skip"): How did I do this past week? Did I miss anything?  
 Alan: The coffee, skip. The coffee. You have to actually  
     turn on the coffee machine in the morning.  
 Skip: I see. Not just scheduling turning it on?  
 Alan: Scheduling it, AND actually turning it on. Both are important.  
 Skip: Ok, I have updated my Tasks & Procedures. Anything else?  
 Alan: Try talking less during meetings. Possibly...try not  
     talking at all. Just to give that a try.  
 Skip: Ok. Anything else?  
 Alan: Hellen said she was most impressed with the microcontroller  
     code you write for her iOT servers. Well done there.  
     Add that to feedback.  
 Skip: Ok. I have updated Feedback & Testing. Anything else?  
 Alan: That's it for me. How about you? Any observations?  
 (check feedback, test  
     Skip: I did notice Johnathan seemed a bit glum this week.  
         Maybe we can all chip in, try to cheer him up. Buy him some  
         flowers, or maybe an artichoke.  
     Alan: Good idea. Let's get him some flowers.  
     Skip: He does like cooking, you know.  
     Alen: Alright, get him an artichoke. Weirdos.  
     Skip: Very well.  
     Alan: Good job, Skip. See you Monday.  
     Skip: Goodbye, Alan.

## Appendix #: Walk-Through Examples and Discussions

### **Part 4: Alone In the Dark with a Bag of Words Outlining & Navigating Object-Relationship Spaces**

Note: This section may be a disastrous bust, with seasoned AI architects face-palming and saying: "Who is this idiot? This isn't how anything works! You can't do it like this!"

What I am trying to do is to illustrate at a very superficial high-level how the AI will need to draw upon various kinds of information. But I am not literally trying to say that it must happen in this rigid 'symbolic' prescribed way. This is just a conceptual illustration to show one (maybe ridiculous) way where the AI works with a variety of objects and types of object-relationships during sections of work.

For example: What is an object "space"? At a very high (fictional) conceptual level, you can imagine the AI taking objects of the same type and putting them into the correct 'spaces' where they can interact. E.g. Two language strings would interact in an NLP-Language space. Two pool-ball math vectors would interact in a math-physics-space. Two schedule items would interact together in

time-schedule space. Now, in reality, the AI isn't a physical wooden cabinet with physically separated system spaces, where you would open the draws and drop the "objects" (which are often not physical objects that you could pick up) inside. STEM traditionally uses the language of spaces. Problem space. Ecology space. The Biosphere. The Chlorosphere. The atmo-sphere! The area "under the curve." There are vector spaces and set theory spaces. So hopefully it is not too unconventional or untenable if I refer to: the space of abstract logic. The space of Natural Language. The space of schedules and time. The space of statistical approximations, etc.

These 'spaces' are types of interactions and the types of objects that interact that way, which can simply be a label or flag or 'attribute' (like a property) that you give to objects to signal that they are of the same compatible interactive type. These are both language strings. These are both logical operators. THese are both Monday meeting schedule tasks. So the AI will be able to proverbially handle both objects in the shared 'space' of system type and type of interactions.

Now I am going to try (try) to illustrate the intersection between these Spaces of object relationships, with the AI dialogues that we just covered. The idea is to show how and AI-bot might use or invoke different Spaces of Object Relationships as it carries out the dialogue.

With some luck, this will also help to illustrate how a more mature and functional Woz-Office-Ai-bot will need to do and navigate and manage much more than a simpler browser chatbot.

Note: The Woz-Test for AI is a reference to this interview,

<https://www.youtube.com/watch?v=MowerqwQR5Y>

where Wozniak describes the ability for an AI to make a cup of coffee as a high benchmark of ability to meet.

internal dialogue and an internal operating system:

How can we step towards our goal?

Het's imagine the AI digests the question or statement input by asking itself questions, to fit the data into an object relationship framework.

Again, this should be the easy part, identifying things is what AI is good at! The problem is that so far there's "no where" to "put" that information.

What are the objects? What are the properties? What does this mean? What is that?

This leads us to the plurality of object-relationship spaces. Notice, we started out with two specific and different examples. You might just have scrutinized how concrete and answerable those were, but those example dialogue were deliberately chosen to illustrate two different object-relationship spaces. Both are concrete, both are computer friendly, but nevertheless they are separate spaces.

So as our AI, Skip, ponders and queries and puts this or that bit of information into the different buckets in object-relationship-space...what does object relationship space look like?? Where are the buckets? So far it seems like Skip is alone in a dark room with a bag of words, grasping around to drop a word here or there in the dark. Let's illuminate the space.

The buckets and functions in object relationship spaces:

A directory of spaces:

- question space
  - library space: reference materials
    - private notes
    - private data
    - external data
  - object property database
  - Plans, Reminders and Notification Space
  - Tasks & Procedures Space(s)
  - Translation/COnversion Space
  - Code Execution Space
  - tautological abstraction space (math, programming)
  - generative 'imagination' space
  - reality space
  - hypothetical/model spaces
    - purely hypothetical spaces
    - alternate/modified-reality spaces
  - Model & Pipeline Sub-Spaces by Model Type:
    - types of DS/ML/AI
  - gamification space
  - quarantine spaces
  - sandbox space
  - ~ "conversion-space"
- Q - self reflection & recursion?  
Q - analogies  
Q categories of types of systems?  
Q system fitness  
Q system defense  
Q: translation?

Walkthrough:

## "everything is an object"

The initial settings for how the OS is configured can be later changed and improved by the AI itself, as everything in the OS is an object in object-space.

Let's run through three examples, based on dialogue between a human (Alan) and an AI using the AI-OS (Skip). Note: The distinction between the AI and the AI-OS may be ambiguous.

Example 1: Talking about Logic & Language

Where SKip and Alan talk about abstract variables.

Example 2: Talking about billiard ball physics

Where Skip and Alan talk about billiard balls bouncing around on a table.

Example 3: Talking about Mixed & Hypothetical Topics

Where Alan asks skip about hypothetical scenarios involving billiards and logic.

Example 4: Self Improvement,

where Skip and Alan talk about themselves and their work.

### **Example 1: Talking about Logic & Language**

Where SKip and Alan talk about abstract variables.

- gets signal
  - sees/ hears Alan
  - identifies Alan
- check schedules & reminders:
- gets signal (from Alan)
- quarantines & checks signal
  - checks raw signal
  - checks ok
  - passes signal to main-OS
- id-scan for objects
  - protocol: greeting \_> return greeting
- Get signal:  
*Alan: If a = b, and b = c, does a = c?*

- id-scan each object for object-space-type
    - abstract space
  - id-scan signal type
    - question: properties:
      - conditional
      - boolean question
      - answer: truth value property of 3rd object
  - translate question into abstract calculation
  - assign truth value to 3rd object
  - generate signal: answer: truth value property of 3rd object
  - Get signal:
  - protocol: farewell \_> return farewell
- + Object Relationship Spaces vs. Symbolic & Subsymbolic AI  
(move to part 1?)

Note: the AI might make mistakes translating the natural-language into an equation, but it should still be able to discuss the 'objects' being discussed, and so recognize that it is being corrected if it makes a mistake.  
(See openAI's section on limitations of chatGPT)  
To some extent (maybe contrary to a Turing test) I am much less interested in the answer being correct, compared with there being some visible object-relationship-space going on.

entity extraction is a work-in-progress, the point is that currently that information is going nowhere for the AI itself.

Note: perhaps like Kahnman and Tversky's system1 vs. system2 thinking in *h.Sapiens*, there a good chance that the first answer to an overall question will be based on a fuzzy estimation process. But if specific questions are asked, deliberate invoking the more logic step detailed system, that could be done.

e.g. Answer to problem: toss it into a model.

closer inspection: what is the first object? what is the second object? what is being asked about the relationship between the two? What was your answer? Was your answer correct?

etc.

Again: the goal is not for the AI to do everything 'symbolically' (or non-intuitively) but to have an awareness of what objects are there, and to be able to deal with them deliberately

## **Example 2: Talking about billiard ball physics**

Where Skip and Alan talk about billiard balls bouncing around on a table.

- gets signal
- quarantines & checks signal
  - sees/ hears Alan
  - identifies Alan
  - checks flags/reminders/notifications
  - checks security
  -
- gets signal (from Alan)
- quarantines & checks signal
  - checks raw signal
  - checks ok
  - passes signal to main-OS
- id-scan for objects
  - protocol: greeting \_> return greeting
- Get signal:

**Q: not micromanaging how exactly the question is broken down.**

- id-scan each object for object-space-type
  - abstract space
- id-scan signal type
  - question: properties:
    - conditional
    - boolean question
    - answer: truth value property of 3rd object
- translate question into abstract calculation
- assign truth value to 3rd object
- generate signal: answer: truth value property of 3rd object
- Get signal:
- protocol: farewell \_> return farewell

### **Example 3: Talking about Mixed & Hypothetical Topics**

Where Alan asks skip about hypothetical scenarios involving billiards and logic.

- gets signal
  - sees/ hears Alan
  - identifies Alan
- gets signal (from Alan)
- quarantines & checks signal
  - checks raw signal
  - checks ok
  - passes signal to main-OS
- id-scan for objects
  - protocol: greeting \_> return greeting

- Get signal:

Q: not micromanaging how exactly the question is broken down.

- id-scan each object for object-space-type
  - abstract space
- id-scan signal type
  - question: properties:
    - conditional
    - boolean question
    - answer: truth value property of 3rd object
- translate question into abstract calculation
- assign truth value to 3rd object
- generate signal: answer: truth value property of 3rd object
- Get signal:
- protocol: farewell \_> return farewell

#### **Example 4: Self Improvement**

where Skip and Alan talk about themselves and their work.

Human-agent("Alan"): Hello, Skip.

```
(check security)
AI-Agent("Skip"): Hello, Alan.
                    How did I do this past week? Did I miss anything?
Alan: The coffee, skip. The coffee. You have to actually
      turn on the coffee machine in the morning.

(check procedures, NLP pipeline)
Skip: I see. Not just scheduling turning it on?
Alan: Scheduling it, AND actually turning it on. Both are important.

(Update procedures, NLP pipeline, feedback, private notes)
```

Skip: Ok, I have updated my Tasks & Procedures. Anything else?

Alan: Try talking less during meetings. Possibly...try not talking at all. Just to give that a try.

(Update procedures, NLP pipeline, feedback, private notes, health, )

Skip: Ok. Anything else?

Alan: Hellen said she was most impressed with the microcontroller code you write for her IoT servers. Well done there.  
Add that to feedback.

(Update procedures, feedback, private notes)

Skip: Ok. I have updated Feedback & Testing. Anything else?

Alan: That's it for me. How about you? Any observations?

(check flags and reminders)

Skip: I did notice Johnathan seemed a bit glum this week.

Maybe we can all chip in, try to cheer him up. Buy him some flowers, or maybe an artichoke.

Alan: Good idea. Let's get him some flowers.

Skip: He does like cooking, you know.

Alan: Alright, get him an artichoke. Weirdos.

Skip: Very well.

Alan: Good job, Skip. See you Monday.

Skip: Goodbye, Alan.

- gets signal
  - sees/ hears Alan
  - identifies Alan
- gets signal (from Alan)
- quarantines & checks signal
  - checks raw signal
  - checks ok
  - passes signal to main-OS
- id-scan for objects
  - protocol: greeting \_ > return greeting
- Get signal:

Q: not micromanaging how exactly the question is broken down.

- id-scan each object for object-space-type
  - abstract space
- id-scan signal type
  - question: properties:
    - conditional
    - boolean question
  - answer: truth value property of 3rd object
- translate question into abstract calculation

- assign truth value to 3rd object
- generate signal: answer: truth value property of 3rd object
- Get signal:
- protocol: farewell \_ > return farewell

?

?If Tom invites Anne to a meeting that begins at 2pm UTC, what time does the meeting begin?

?

- (As to how the AI might search through these spaces, here is one idea.)  
 As an initial brute-force example (with later fine-tuning built into it):
- The AI has a current "state" of objects.
  - If a signal comes in (I say 'in' but it might be a self-scheduled notification or something):
    - by default run through each space, especially procedures, to see what applies. Most won't, but some will.
    - Security & Quarantine & Health checks will apply to all incoming signals.
    - Then probably run a first identifier sweep for low-level files.
    - Then if it is a media file, or unknown type file, or .txt file, follow a procedure.
    - Security & Quarantine & Health checks will apply again to contents of files.
    - If it's a media file about an 'object' in the office, you might check catalogues of types of systems to see how it works. (e.g. run an ID model and add an attribute-flag to that object)
    - disposable, long term
    - questions: if you don't know what it is or what to do with it, ask someone.
  - This may be a quick and optimized search in a deployed system, being able to search (e.g. in an efficient dictionary) and only hit what applies
  - and it might be a search that runs several times on several sweeps (several layers?)
- eg.g
1. first sweep, unopened signal, go to security
  1. security check 1: ok open signal
  3. security check content: ok
  4. content secure, identify content type
  5. sweep for signal type (e.g. image file)
    6. AIML model id file: file is birthday picture

7. look for birthday image notes, e.g. in tasks and notifications  
this was an expected file: see procedure,
8. follow tasks scheduled procedure: requested: image forwarded to team member

Note: most likely this kind of crude brute-force workflow will be optimized by the AI itself as it gets experience with handing various kinds of signals, likely no need to run every signal through every check option. But there aren't many check-options here yet so this may be a viable place to start.

A note on object oriented programming:

- not a fan
- What has the OOP bro-coder cartel produced since the 1970's? A dumpster fire. Crisis would be a kind way to describe the state of computer science in ~2020.

More simple example:

A bag contains 10 marbles.

A jar contains 50 marbles.

Which contains more marbles, the bag or the jar?

or

A bag contains 10 marbles.

A jar contains 50 marbles.

Does the jar contain more marbles than the bag?

Less Simple example:

If I have two containers, a bag-like container with more than a dozen objects in it, and a jar-container with merely a pair of things in it. Which container contains more objects?



Jar > bag.

chat.openai.com

Also note: chatGPT dynamically named the conversation accurately after the topic being talked about. " Jar > bag. "

A bag contains 10 marbles.

A jar contains 50 marbles.

Which contains more marbles, the bag or the jar?

The jar contains more marbles than the bag, as it has 50 marbles compared to the bag's 10 marbles.

A bag contains 10 marbles.

A jar contains 50 marbles.

Does the jar contain more marbles than the bag?

Yes, the jar contains more marbles than the bag. Specifically, the jar contains 50 marbles, while the bag contains only 10 marbles.

If I have two containers, a bag-like container with more than a dozen objects in it, and a jar-container with merely a pair of things in it. Which container contains more objects?

Based on the information provided, the bag-like container contains more objects than the jar-container. Even though the exact number of objects in the bag is not specified, we know that it contains more than a dozen (i.e., more than 12) objects. On the other hand, the jar-container contains only a pair of things, which means it has two objects. Therefore, the bag-like container contains more objects than the jar-container.

kiku ai: (classic!)

A bag contains 10 marbles.

A jar contains 50 marbles.

Does the jar contain more marbles than the bag?

What is giant sand?

How much would something like that cost?

Surely, any jar would have a lid?

Wolfram AI:

A bag contains 10 marbles.

A jar contains 50 marbles.

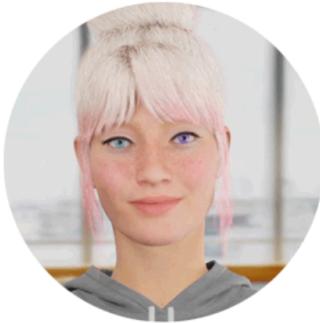
Does the jar contain more marbles than the bag?

Interpreting as: contains

Input interpretation

contains (English word)

@kuki\_ai



Discord

Video Chat

© ICONIQ 2025 | v18.1-230126



Sure. Ask me anything you like.

If I have two containers, a bag-like container with more than a dozen objects in it, and a jar-container with merely a pair of things in it. Which container contains more objects?

How much does it cost? thumb up thumb down



The blue one, I think.

Type a message



Voice Chat



Gifts



Games



Readings

@kuki\_ai



### Heartache Of The Telephone Industry

□□□□□□  
□□□□□□  
□□□□□□  
□□□□□□  
□□□□□□  
□□□□□□  
□□□□□□



Discord

Video Chat

About

A bag contains 10 marbles. A jar contains 50 marbles.  
Does the jar contain more marbles than the bag?



What is giant sand?

How much would something like that cost?

Surely, any jar would have a lid?

While ChatGPT is significantly closer to having the abilities in question:

1. identify objects
2. give **attributes** to objects (or modify, etc.)
3. predict how objects will interact
4. plan events based on objects
5. take actions based on plans
6. communicate about the above

We will be talking about general and specific skill sets. ChatGPT does appear to consistently be able to identify objects within discussion, which is a big important step that no other AI I have tested has broadly and consistently been able to do.

The goal is to move from a chatbot or other single-function AI, to an AI that can perform a variety of specific modular tasks within Object Relationship Spaces.

There often is still the gap (though much smaller with chatGPT) between a 'spew everything' answer which arguably somewhere includes relevant information, and being able to focus actions and dialogue on specific objects and **attributes** of objects.

The goal here is not to claim that current AI can never perform these tasks.

The goal here is not to get bogged down in terms or goals that cannot be clearly defined, terms such as: "intelligence" or "understanding."

The goal is to try to use the concept of Object Relationship Spaces to evaluate, discuss, plan, and improve the functions of AI in clearly definable and testable, and modularly-expandable, ways.

Human: What is this?  
AI: That is a white ball.  
Human: And what is this?  
AI: That is a black billiards ball.  
Human: Are they both balls?  
AI: ...I found this on the web: a billiard is a kind of bird...  
[FAIL]

or

Human: What does this video show?  
AI: That is a video of a moving white ball.  
Human: Will the moving white ball hit anything?  
AI: ...I found this on the web: Moving to White Ball Wyoming?  
Best Ranches to Buy! on buy-your-ranch-now.com...  
[FAIL]

When an AI correctly identifies a cat picture does the AI realize what it has done? Does the AI know that a cat is a 'thing' or an object in a space where objects have properties and interact with other objects? Does the AI realize that it has just identified an object, so there that is (effectively) then both the object which has been identified and the identification of that object (where the identification is also an object of some type)? Generally the answer to all these questions, as of 2023, is no.

The standard explanation of the limits of AI by nearly all sources goes something like this: You ask an AI to identify what is in a picture, and yes it correctly labels the picture as "cat," but the AI has no idea or thought or understanding about what it just did. It doesn't know what a cat is, or what a picture is, or what an identification is, or that you asked it something, it is just a pre-canned automatic process that spits out that label when you enter that input.

And then there is usually some filler explanation that goes something this: In order to have real, true, pure, understanding and intelligence, you need to

have a knowledge base, and you need to have general intelligence, and you need to be able to flexibly apply and transfer learning from one area to another, etc. etc.

Before ChatGPT, this standard explanation was easier to say because the granularity of the language responses was so rough/crude/poor that you could not ask the AI about any objects or relationships within what you were talking about. That is now different: ChatGPT can usually do this, yet people still say the same undefinable thing about ChatGPT "But the AI doesn't understand what it is saying." A more narrow but similar counter-example, while it is much less consistent or flexible, when you give Wolfram Alpha exactly the kind of equation it was designed to compute, and you are then able to ask detailed questions about each part and result of the equation: each object and its relationship to each other object with all the mechanics explained and illustrated. In that admittedly extremely narrow example, it is much more difficult to say that Wolfram Alpha does not understand logical equations. Probably people would pivot and say: Well it can't do anything else!

[This consensus on 2020 AI's inability is problematic for 3 reasons:

1. It is currently being contracted by ChatGPT. ["Current AI will never be able to write or evaluate computer code. That would require true human understanding!" Sounds like what people said about chess, doesn't it?]
2. It is based on a problematic comparison with fantastical exaggeration of human intelligence and learning. ["Humans learn and generalize to everything instantly everywhere...more hyperbole, etc."]
3. It is based on terms and goals that either have not been or cannot be defined. ["Understanding," "general intelligence"]

It may be that we are so lost that we don't even know what AI isn't doing right, in part because we know so little about ourselves and our own so-called 'general' 'intelligence.' So I want to take a different approach from most of the largely excellent books out there that try to explain the limitations of AI in terms of 'generalization' and 'understanding.' On the surface, yes, of course, I agree, I do not think that AI is secretly micro-managing layers of hidden intelligence. AI now is often very shallow and fragile, and if the only point of those books is to explain to scared consumers of conspiracy-loving people who think their toasters are going to literally take over the world tomorrow then I have no complaint with that. My goal here is to try to make a serious road map of performance requirements to make AI better. I'm not interested at all in spending one picosecond worrying about defining terms like 'understanding' or 'intelligence.'

This is a proposal to make some steps beyond this current "mindless" state of AI, to try and give AI some ability to do something with the information it produces. How well it will work...well version 1.0 probably won't work well at

all! But worth a try. Let's follow the logic of our situation to claw our way to something that works incrementally and sustainably better.

We will look at two specific abilities that an AI could have, or two levels of related abilities:

1. Can an AI apply an identification to the same object that it just identified? Goal 1: Objects in Object-Relationship Spaces have properties.
2. Can the AI model (e.g. predict, score, recommend, generate, etc.) how two objects might relate to each other (including interaction with each other)? Goal 2: Objects in Object-Relationship Spaces interact with other objects.

For example, this is how a (hypothetical) conversion with an AI might go in 2023.

Human: What is this?  
AI: That is a white ball.  
Human: Correct! What color is that white ball?  
AI: ...I found this on the web: color is a relative social construct...  
[FAIL]

Conclusion to Part 1:

So far this narrative has been intended as a brief and casual outline and introduction intended for all audiences on the topic of AI-in-2023's thin-to-zero ability to navigate Object-Relationship Spaces.

"Understanding" and the turing test:

My narrative may be walking the line between topics. If this is yet another dichotomy between 'understanding' or 'intelligence through understanding' vs. 'intelligence without understanding', vs. a kind of Turing test aspect where a performance test should suffice: Personally I don't think 'understanding' is clearly defined here at all, and I'm not arguing for a specific approach to being able to handle tasks involving object-relationship space. And part of the problem is that performance is what is lacking. If AI was so good at faking object-relationship-space, I wouldn't be writing this.

My aim is to use specific framing about object-relationship spaces to more concretely describe what things current AI cannot do, broken down step by step until they are things that we can figure out how to do.

When it comes to talking about what AI can't do, many books about AI (see below) focus more generally and vaguely on concepts like 'generalization' in very strange comparisons to people. I have a background as an educator, so however bad a teacher I was, I have spent lots of time either succeeding or failing to help people overcome innumerable obstacles to learning. Human learning for most people is an agonizing process that takes many years, often decades. And that's just learning. When it comes to project management people are like some kind of nightmare device designed by a supervillain to destroy the universe. Seriously. People fight absolutely every part of trying to get anything done. They hate schedules, they hate roles, they hate talking, they hate listening, they hate learning, they hate goals, they hate deadlines, they hate stakeholders, they hate transparency, they had planning oh my how they hate planning, they hate showing up to work, they hate their kids, they hate their parents, they hate themselves, they hate their own vacations, but they do love their pets, some of them at least. And then all these books go on talking about AI saying the same incredible things: humans have perfect instantaneous self correcting infinitely flexible absolutely general intelligence! Are you kidding me? Seriously, have you met a human being? In the introduction to Daniel Khanman's "Thinking Fast and Slow" (An amazing and widely read book which shockingly it seems like no data scientist has ever read.) he talks about back the 1970's when, somehow just after WWII almost destroyed the planet through epic bad decision making, psychology-academia decided that everyone has a perfect ability to model and solve all problems including absolutely perfect intuition about statistics: that's it, case closed, topic is closed for discussion, that's the truth, just move on. So Khanman and Tversky spent decades re-introducing the concept that people can make mistakes to endless vicious trolling hostility from an insane paradigm that assumed that no one ever makes bad decisions (a little irony there?). T&K came up with innumerable studies painstakingly breaking down step by step how people's intelligence, understanding, and problem solving abilities are hot garbage. Eventually Kahnman won a Nobel for the research after it was used to found several new disciplines of study (Tversky had died of cancer by that time). So let's get serious about this, folks. People do not have general intelligence. People do not have automatic learning. People do not have automatic transference from one nugget of learning in one area to infinite fantasy multiverse rainbows everywhere.

In a sense there may be another level here to the old symbolic vs. sub-symbolic AI debate around this set of Object Relationship tasks. History note: In the pantheon of horrible naming practices, the name 'symbolic vs. sub-symbolic' was (heaven knows why) given to two schools of thought for approaches to AI. Step by step direct logic was called "symbolic" (?), and deep-learning type AI was called "sub-symbolic" (because hey, why use language that makes sense, right?). My main interest is in outlining the goals and tasks (of Object Relationship Spaces) themselves, and perhaps ideally letting each AI find whatever the best solution is that arises over time. If someone is extremely anti-symbolic, in theory you could make an ensemble of parts that use sub-symbolic solutions to meet all of the performance requirements outlined here.

You could simply have:

NN 1 answer the question (or answer the signal, maybe it's not a 'question')

NN 2 break question into objects

NN 3 connect bits of answer to objects

The depth of the specific Object Relationship Tree (if that's an ok phrase) can depend entirely on the context. For example, without a context, there would be a ridiculously infinite way of breaking everything down in every way into every possible set of parts, physically and otherwise, and having a gazillion quarks isn't going to do much except use more resources than you have to spare.

See example of billiards questions:

What is this?

a billiards table

What is on the table?

Balls

How many balls are on the table?

7

Are all the balls the same color?

no

Is there a neon-green ball on the table?

no

Have you seen a ball before?

(look in object database)

Are balls round?

Yes, they are.

Comparing AI to Human 'Intelligence'

### **Low-bar enlightenment: a human halting problem & 'Agile'**

While the topic of 'mindfulness' or other heady discussions of the nature of consciousness and reality may quickly slide down the slippery slope of being no longer useful to a computer science project, I would like to make a very brief and concretely-defined observation: While the main comparison used when attempting to frame and define artificial intelligence (or whatever synonym or set of terms a given author prefers) is between non-human AI and Human H.Sapiens. Nearly universally the 'intelligence' of h.Sapiens is vaguely and hyperbolically described as absolutely fast generalizable flexible, etc. Yet there is a huge fundamental (and perhaps ironically problematic for completing AI researching projects in academia or business) with human intelligence (one of many problems): h.sapiens has no awareness of being trapped in cycles of system collapse, making the same mistakes over and over without learning, making any effort to end cycles of failure, showing no interest in the concept of cycles of failure, not being able to see the cycles of failure. As just one example of how we need to nuance our dichotomy between AI and perfect-human-intelligence, the ability to see that a process is failing to do something about it is:

- a crucial part of AI

- a crucial part of h.sapens future survival
- a crucial part of how h.spenns and AI will work together

methods for consensus  
doing STEM studies  
hypothesis testing  
policy  
policy and system collapse  
elections and consensus  
logging actions  
the question of bad-actors  
incentives  
the question of network access  
cyber security AI monitor  
'internal consensus'  
Ecosystems of AI components on possibly large scales  
e.g. distributed weather monitoring  
distributed earthquake monitoring

the question of reproduction

people should do their own specific tests where possible  
- it is very strange for a person to write about a public facing AI chat-bot, where all they do is quote a far away expert saying

The magic flying car question: it's very likely the wrong question.  
At what point does this have X?

The only context I can think of is kind of fuzzy sales-marketing context where you are trying to model what customer's want in bazaar aggregate, and so you will get these weird overlapping generations:

- people want a comfortable, grandmotherly, exciting, james-bond, flying-broomstick, with family comfort, like oatmeal and chicken-soup, but with lots of chocolate cake, which you could bring to work and take to church, and good for bird-watching, but also enjoy on the sofa at 3am, and get's really good gas milage! What should we call the sweet spot in all that?

People rolling out a family sedan are going to have to make that kind of hodgepodge, and they might nickname it 'magic flying car' but in a sense even there they are at least trying to refer to specific (if sometimes contradictory things (what the 4 year old boy wants and what his mom wants and what the family dog wants and what the grandfather wants, etc.) which you could image as...back to AI and machine learning....vector in those directions in a multi-dimensional space. So even here there is a measure-able or aim-able goal.

As they are not, terms like 'intelligence' and 'understanding' and 'general intelligence' and 'human intelligence' simply have no clear meaning, and are sadly looking to be in the same category as gestalts from history that just utterly failed to connect with reality. Perhaps like the 'precious bodily fluids' of the comedy-horror-satire Dr.Strangelove, blind as we are in any given moment as to where we are and how to weight short-term and long term value, obsessing over things that do not exist and using those to try to triangulate goals, states, conditions, outcomes, etc., likely will not propel the human story in directions that in retrospect (e.g. from seven generations in the future) will be considered to have been a good idea. E.g. Mixing fiction and reality usually leads to system-collapse. Like the house always wins in a casino, but perhaps worse. If you're throwing coins away where there isn't even a slot machine, your odds are significantly worse than being forced to gamble with least-worst options for your own survival. As a species that is very likely simply too stupid to survive (to unable to see, to unable to learn)

#### AI, Compassion, Empathy, & Ethics:

interesting to think of chatGTP as being 'visually impaired'

IQ test

testing 3D space

humor, compassion,

violence, radicalization, paranoid-conspiracizing, blind cycles of self-destructive and community-destructive behavior.

...

#### The cautionary example of Chess:

Chess is junkfood, it is a sport-game that has evolved to attract some human minds because it is attractive, not because it has substance. Chess does not correlate or integrate with anything good that is, significant. And chess does correlate with the worst biases of bro-culture coding and politically biased psudo-science.

For example, the two main pioneers that made computer chess possible: Alan Turning and Jon Von Neuman, both super-genouses who created muliple new fields etc etc., could write whole

programs in assembly language as if were python, were both famously horrible at playing chess. Chess is a mind-illusion, a

For example: Bro-coder tech culture thrives on bullying 'hard core' gaming and 'hard core puzzles' which somehow excite and motivate bullies to dominate other people, but which serve little or no practical purpose. Yet in the minds of those bullies, they conflate their own bad habits, addictions, and abusive behavior with rationalizations about technology helping society. This is a dangerous combination. They see themselves as superior coders and puzzle solvers and the kings of STEM space, whereas in reality they are petty delusional drug-addicted tyrants torturing the local peasants while Rome burns around them. Since the 1970's, the socio-paths who dominated computer programming left the future a legacy of catastrophic neglect and incompetance, while destroying the lives and careers of 'weak' people for their own personal amusement. This is a nightmare scenario, and we are living in the results of that nightmare.

The founders of computers worked from the 30's to the 50's to put together the foundation:

- Alan Turing
- Jon Von Neuman
- Claude Shannon
- Alanzo Church
- Grace Hopper

And many other less famous but no less important people.

The first generation of mature-ish computers created the amazing combination of POSIX, Bash & C in the 1960's, which we are surviving on today. But this is like waking up on a geneation-ship built generations ago by long dead geniuses and finding that no one has even done basic maintenance or upgrades for fifty years and there are constant explosions erupting all over the ship (while war-lords wage battles in the light of the burning piles of garbage).

There was also a brief 'first internet' period in the late 1960's based on mainframe terminals which most people seem to have completely forgotten about (see: A People's History of Computing (a book which is sadly overshadowed by political ranting, but which is very worth reading))

And in 2023, people are working on the 2nd generation low level language, called "Rust," and are barely starting to work on a second generation of OS systems (which is not to say that POSIX is in the dangerous position of C, which is radioactively memory-unsafe and should essentially be a museum piece and never used for publicly available software; though in retrospect maybe people will say that was the case that 2023 OS kernels were beyond fixing).

This means there is a 50-year gap. What was happening between 1971 when Nixon closed the gold window and 2022 when Linux started talking about introducing provably memory safe code into a POSIX OS kernel? If we survive it will be interesting to see what historians have to say.

Does this mean that chess cannot be used for computer experiments? No.

Can chess be used to illustrate things. Yes, of course.  
But be careful with chess, and think carefully about what you think chess is.

We should look very carefully at the details of how very wrong people were about their explanations and predictions about AI and computer chess playing.

...

(a definable alternative to "general intelligence")

In perhaps a classic example of "Garbage in Garbage Out" (or, as I think of it: you cannot make a good thing out of bad parts.)

If you start with one indefinable thing (with is something that does mean anything), and use that to make multiple undefinable things, you will have started with one meaningless thing and used that to make multiple meaningless things, which has been a waste of time that has not accomplished anything meaningful.

People start with the undefined and meaningless term "intelligence" and proceed to turn that into:

1. General Intelligence
2. General Human Intelligence
3. General Machine Intelligence

The painful narrative goes something like this: AI isn't great yet because it lacks General intelligence. What is General Intelligence? AI doesn't have "General Intelligence" because it isn't a person, and people have "General Human Intelligence". Aside from all actual know studies of human psychology and behavior, we computerscientists are going to spontaneously declare that humans have this all powerful black box ability to do anything instantly and that's called: general human intelligence. Machines will never have "General human intelligence," because they are not human beings! But some day they might have "General Machine Intelligence," but no one knows what that is, or when it might happen, or how.

This whole narrative is completely meaningless and counterproductive. It was meaningless from beginign to use the word 'intelligence', it was meaingless all the way though, and the sum total of it is just as useless and meaningless. Garbage in, garbage out, it is meaningless.

I started working on Object Relationship Spaces before chatGPT, but I think chatGPT has moved everything ahead.

1. There was a lot of pontificating about how embedding-type AI could not do what chatGPT then did do because of a completely undefined "lack of general intelligence."

I originally created (or re-created) the context of object-relationship spaces to more clearly map out:

1. What the goals for specific AI abilities are.
  2. What the tests should be, and how to define evidence.
  3. What we are claiming AI can do.
  4. What we are claiming AI cannot do.
  5. How we define problems, failures, and obstacles.
  6. What we expect and plan to be requirements to meet goals and overcome problems.
- etc.

ChatGPT I think has thrown a bit of a spanner into the works, taking a situation that was already confusing and confused, and adding many more not-yet-understood elements to it. Abilities that people said were not there, apparently were. Obstacles that people said were there, apparently were not there. Goals that people said were not being aimed for, either were or were achieved by accident. And the diversity of abilities in the overall set of abilities that we already did not understand just got a lot bigger. With the result that people kept blithely saying exactly the same nonsense things about chatGPT that they said about earlier completely different forms of AI, making those old cliche comments more diversely wrong in new and exciting ways.

If someone wants to use Object Relationship Spaces as a way of looking at or defining 'intelligence' or 'general intelligence' that is their choice.

A similar cautionary tale here may be the history of the IQ test, or the "intelligence quotient" test. I recommend Stephan J. Gould's books the mis-measure-of-man.

(And this story might rhyme with the computer-science dark-age story where between 1970 and 2020, progress was extremely hampered by bullies wasting time pleasuring themselves with logic games and dominance games that they personally enjoyed).

What today is called the IQ test, and which is sometimes used to measure how intelligent a person is (as though measuring the static height of a table top), was developed very specifically only for use in identifying people who were not getting enough of a given educational resource so that they could be given sufficient educational resources to not fall behind in that area as they continually developed, in a context of continual growth and learning by all people. It was very specifically never to be used to pretend to be able to measure the total "intelligence" of a person, especially in a static scenario where people are finished learning and developing. And what happened? Sociopathic men turned that into a set of their favorite logic puzzles that they enjoyed using to bully people they perceived as being weaker. Not exactly a tool used as intended by the science and people that created the tool.

As a future end goal, if we can eventually come up with a meaningful set of frameworks around 'intelligence' and 'wisdom' and 'consciousness' that would be great. In the meantime, if people want to philosophize and speculate and research and make art about the mind, that's great (my understanding is the science fiction literature of the 30's to 60's had a very positive simulating effect during the education of later generations). At this time (2023) we know so little about the

mind and consciousness (from the brain, to mindfulness compassion, to entheogens, to much else) that I have no grounds to speculate what a future hard science would be, or even if we are smart enough to survive long enough to achieve such a science. We should research, we should work hard. We should be extremely vigilant against bullies wasting time pleasuring themselves with logic games and dominance games that they personally enjoy.

...

### Projects: the context is projects

agile projects and the hypothetico deductive method are perhaps the two most important cutting edge breakthroughs of science and STEM that most people don't even know exist.

The reason why I am using these terms is that the spaces (be they literal or proverbial) of mind, language, intelligence, consciousness, projects and project management, information, etc., have historically been so difficult to describe that even today, to be honest, people have simply preferred to not bother to study them. There is no intelligence-ology, or consciousness-ology, or learning-ology, or education-ology. Our species is content to say that we are the center of the universe because we alone can do...this stuff...and that should be good enough for you, move on, nothing more to see here. In the appendices below there is a critique of a cross section of AI-literature to date and it's concepts and lexicons. But it is in an appendix because I don't want the red herrings and undefinable terms of most books to be a distraction here. The goal here is to focus on clearly defining how we can expand the function of AI in practical ways, maybe that has nothing to do with XYZ undefinable term...that's totally fine. We are here now to focus on clearly defining how we can expand the function of AI in practical ways.

In a sense, the term 'object' and 'object space' could be so broad as to be problematic. Anything can be an object, and the 'spaces' for anything and everything are too big and vague to bother with. But context is key: And the context here, is what AI can do with things that it comes up with. And in that context, and 'object' and the relationship between objects, is very important, and also very clear and practical.

We need something like a notion of networked-intelligence.

(Note, I'm trying to describe, not opaquely name.)

Multi-faceted intelligence  
multi-channel intelligence  
multi-signal intelligence  
multi-space intelligence  
Multi-Stuff Intelligence (stuffy intelligence?)  
multi-things intelligence

It is not possible to clearly discuss AI without being hampered by intimidation, politics, threats, etc. By which I mean, we are required to make assumptions that are demonstrably false at the onset just to have a discussion of AI. The very wrong and very bad background assumption we are required to make to be allowed by society to discuss AI is that:

- people won, people did everything, we've discovered everything, we've done everything, learning everything, made everything, solved everything, and NOW we turn our attention to making AI (perhaps to replicate our omniscient omnipotent in pure technological form). That is a completely false narrative and way of looking at the world and as such pursuing that will go badly.
- We haven't finished studying the world, we haven't finished studying ourselves, we have no finished frame of reference, and yet we are working on what we call 'AI,' and we routinely undermine our ability to describe it and navigate the space of what it should be or can be. (Not to mention the very poor state of computer-science and data-science in terms of being a birds-nest mind field of horribly confusing terminology and neglected lexicons which from book to book no two people seem to use or understand the same way.)

There is no absolute sequence for this,

I also think that increased research into biology (including non-human nervous systems and non-central-nervous system decision-making including plants, hybrid-ai-biology, and into h.sapiens consciousness are advisable. Not to mention integrating digital information technology into DNA-based & RNA based or integrated systems. It is painfully ironic that after a modest boost in AI performance leading to another AI-summer, authors of books critiquing AI are finding themselves needing to compare current AI to humans, yet being in a desert of non-research into human intelligence where virtually no one is seriously studying human intelligence because has long been a 'career limiting decision.' How about this, how about we start learning more about ourselves and start solving our problems.

**Game-ification:**

'Generalization and Transference': The Fallacy of General-Human-Intelligence

- Input Output Metrics
- Agile

Unfortunately, along with not studying mind-consciousness-intelligence, only barely studying project-management, there is a lack of study of what I try to describe at categories of types of systems. That is a whole separate topic(applied philosophy and methodology of science), and I don't want to get side tracked, but I think is likely relevant.

'Science' has been extremely slow to either generalize STEM (as in the infamous problem both in the US and the UK after WWII during their projects to develop the first digital computers, that no one could understand the concept of engineers and mathematicians working together on the same project...so they didn't...so the projects didn't get done), or to not-generalize or make singular types of processes. How do we put together all these things at once?

- people assume there is only one process for everything
- people assume there isn't any commonality between professions to work together (like a single process of how things work)
- people don't study how they work

Yes, altogether, people are not looking too smart.

## **Other notes:**

### **Notes Section:**

-

- Bot Behaviors regarding active Objects in a Networked AI world, described by Object-Relationship Spaces

A Birthday Story

A schedule Object:

The story of a schedule object:

Part 1: Functionality

On Monday morning at 5:30 Alan sends an SMS text message to Skip (the woz-AI office bot at their R&D lab on the southern coast) saying:

"crap, Violet's birthday is tomorrow, totally forgot, morning meetings are canceled, instead a birthday party, probably at noon, ouch that was hot, sorry I'm talking to my phone and just spilled tea all over my foot. I hate wasting tea. See at work skip, thanks a million."

At 6:07 am skip gets a voice mail message from Johnathan:

"Hey skip, sorry my phone isn't working so I'm using Angela's phone, and we're in a cab. This is Johnathan from work. I hear we're planning a party for violet, let's plan to use the main lab room at 11:30am. Try to schedule that if you can. And it would be great to have hot coffee ready this morning if you can do that. Thanks, Skip."

And the story goes on.

1. The schedule keeps changing,
2. the places keep changing, people send mixed messages.
3. Different people do different tasks.
4. Task assignments change
5. Entities need to be identified in various media:
  - website text
  - sms text
  - voice mail
  - video
  - photos
  - gifs
6. materials need to be created:
  - gif birthday sms with text
7. audio messages need to be converted to text and send between people
8. text messages need to be translate to other languages and sent between people
9. pdf files need to be converted to text and image files
10. all people in roles need to be updated about updates to plans and roles, schedules and tasks
11. the completion of tasks have to be checked and reported on (has the coffee maker been turned on, has the card been sent, is there still tea in the tin, is Robin on her way from the airport by rideshare, did someone order flowers, did the address get updated, did the payment method get updated, etc.)
- 12.

The AI needs to convey messages from one person to another, update the overall plan (constantly), schedule and reschedule events, etc. and all with other participants. And everyone wants hot coffee even though that's not directly related to the main topic (except Alan, who only drinks tea). And eventually something actually gets cobbled together: and that's how every day works in

reality. And there should be project review and feedback to improve for the next time.

The AI is going to have to really deal with objects in this networked object space and know how to deal with them. An equation is not the same as a scheduled invitation. A schedule conflict is not the same as hot tea. And these 'objects' have to move around in a real external reality and various parts of AI functionality: input, processing, output, transformations, predictions, recommendations, records, reminders, feedback, reviews, roles, versions, with many different people (and likely AI-agents) participating. There will be pictures and audio and messages and texts and videos and text strings and numbers, and it all has to somehow flow in big networked workflow space, of which the AI is one part, and where the AI likely has many parts.

#### Part 2: Personality & Work-with-able-ness.

Just as people are not intelligent but they do have personalities and hopefully those are personalities we can work with, so of AI. It may be more important that AI is something that can be sustainably worked-in in the long term rather than being highly pedantically correct but not good to work with. So this may be another system requirement, the AI does not need to be operationally perfect (no person is!) but it must be work-with-able...enough.

You have to be able to talk about discrete objects across levels and layers. ChatGPT has made huge progress with being able to get Level one right nearly all the time: if you ask it about objects within what you are currently talking about, it has the 'granularity' to talk about them. But as soon as you go the next level, and you start talking about that conversation, the AI reverts to puking random descriptive information.

#### [The Tea Story](#)

##### [Flat version:](#)

Linda asks Tom for some tea. What should Tom give to Linda?

##### [Networked-Objects Version:](#)

1: Linda orders tea, so Linda's order is: 'English Breakfast.' What did Linda order?

2: Tom took Linda's order. Question: Who took Linda's order?

3: Tom gave Linda's order to the barista. Who did Tom give Linda's order to?

4: Who has Linda's order after the order is given to the Barista?

5: When the barista looks at Linda's order, what does the barista find the order to be? What is the order that the barista sees?

6: When the barista fills that order, what should the barista make?

7: What should the barista give to Tom?

8: What should Tom give to Linda?

In this scenario (which is simple and has no ambiguity or changes or problems or contradictions or even schedule items, etc.) in theory you could try to make an 'all in one' AI that vaguely uses embedding vectors to correctly answer questions about objects in question, but can this work in real life?

E.g. in real life, let's say the barista is an AI! Or that Tom is an AI. Won't we need some kind of actual database of objects and object space to carry out projects with roles and goals in object-relationship space?

"I've got it all in my head."

Again, making connections between AI and people (often people not being intelligent), a classic and horrifying line that is familiar to people who struggle trying to 'herd cats' in project management, is when a member of a team refuses to make plans, make schedules, make documentation, communicate, manage versions, manage roles, simply saying: "I've got it all in my head."

There are no parts, there are no objects, there is no database, there is no granularity. Like a crude AI, crude people often try to handle elaborate situations with many parts as single undifferentiated blobs, puking out splatters of language strings which at best contain some using bits.

While I do not think there can or should be one-way that this should be done, I think sustainable solutions must be non-blobs. The information must be externalized into some kind of searchable, changeable, communicable, environment, an operating system.

We seem to default to referring to single models as 'an AI' while we perhaps should be requiring AI to be both a network of different parts and also an open-network of object-information about 'current projects.'

For Skip, the Woz-AI office bot. If someone wanted to know how many cups of tea Skip made during the past week, and how many cups of tea Skip was scheduled to and asked to make, that information should exist in an open project-information framework of some kind (whatever works best for whatever different application, maybe 'database' is a general enough term, maybe not. People are very touchy about semantics. To some people 'database' is a collection of information, to others it is more specific. I will default to being as clearly descriptive as I can, so I will just say a collection of information).

I do not think scalable solution, involving many participants, roles, interconnection projects, supply chains, etc. could function if each AI was a

'blob' where the only way to get at any detailed project information was to ask it and receive a messy puke of language that contained some of the answers.

"I've got it all in my head."

For managing a project

- People are not intelligent, but they do have 'personalities'  
(not quite the same topic, but the personality of Kiku vs. the personality of chatGPT)

signal, self, interpretation:

- signal directed at self-object:
  - the fallacy of general human intelligence
  - asking questions to fill-in missing information.
  - objects and object oriented programming
  - Francois Challet: right and left hemisphere
  - Mystery of Kahneman Tversky System1 System2 vs. (perhaps a more proverbial) right and left hemisphere.
  - Melanie Mitchel
  - The assumption of 'pure AI' is likely as inaccurate as the assumption of 'perfect pure nature h.sapien's intelligence'
    - AI biological hybrids
  - F Chalet makes the excellent but rarely made point that the available hardware has a disproportionately large impact on what algorithms get used (those most optimized for that hardware). There are perfectly good, in some

ways better, AI technologies that not yet being used because of the relative prices of different kinds of hardware.

- Participation
- Generalized STEM
- Definition Behavior Studies
- 
- self management of resources (e.g. energy use)

Goals:

spaces:

- style, essence, sentiment spaces

These are examples for the purpose of illustration intended for a human audience, but for practical reality, especially given that AI is already too at identifying and categorizing and classifying things, it very well may be best to let the AI create whatever classifications and buckets of objects it wants to keep track of things, even humans do not 'understand' what those buckets are and how they work. (Though if the system overall is working, humans should be entirely able to talk to the AI about those buckets as objects.)

Not Goals:

Through the wonderfully fascinating and oddly poorly documented history of AI there have been many experiments with varying agendas.

Some experiments (or simply books and papers) have focused entirely on Analogy-making.

- making analogies
- playing chess
- winning a turing test (I do like turing tests, but this isn't a goal here)
- Making people feel comfortable (see Eliza)
- Trying to act just a like a h.sapiens human.
- common sense: having an extensive built-in knowledge base to know everything about everything

Goals of Object Relationship Spaces:

- Can the AI do something with the information it generates? (e.g. an identification)

- Can the AI track and attribute properties to specific 'objects.'

e.g. If Skip asks the person a question:

e.g.

Skip: What time is it?

Alan: It's about noon. Does that answer your question?

Skip: Not exactly. Can you be more specific?

Alan: Let's see, it's 11:49am.

Skip: Thank you very much. I have a meeting at five of.

Alan: Don't mention it.

This is a dialogue that we should be able to have with an AI, as it deals only with a completely small discrete finite space of things. The AI (Skip) asked a question, so it can have identified that as a question-object in this given object-space.\* So when Alan says "your question" it should not be unreasonable that Skip(AI) can identify the only question in the question space as the question-object in what Alan says. Again: existing AI is already very good at identifying things, that's not the problem. The problem is that existing AI has nowhere to put, and not it can do with, that identification.

A timeline, a schedule, can be either a fuzzy sequence or a very math-specific number-line. Perhaps like multiple-bridge and traveling salesman, you can contrive a confounding schedule problem (likely offices do so routinely), but schedules don't have to be a mess. There is one event on Skip's(AI) schedule: a meeting at 11:55am.

Note: And a built in part of this, is that if an ambiguity arises, such as (which question?) that is a perfect opportunity for the AI to ask that question! If there were two, or three questions, and it was not clear which was being referred to, the AI can (and should) simply ask.

Again, the point isn't that that is a perfect system. Maybe the AI will try to ascribe something to an object but pick the wrong one. Maybe it will assume that Alan was asking about some other question. The point is that there must be discussions about these objects and their relationships. And again, the sad gold standard (or rusty old bridge standard?) of h.sapiens is appallingly bad at this kind of discussion about schedules. People usually make huge mistakes with schedules, so the bar isn't all that high here. (Perhaps like the relative safety of self-driving cars. Fatalities from human driving are so obscenely high that it might be a challenge to try to design an AI that could cause more fatalities than human's indeterminate incompetence and malice. In fact, though being 'realistic' isn't a goal, the AI would probably be much more 'like-able' and 'human-like' if it showed a natural incompetence with schedules:

Alan: Oh, Skip. Are you coming to the meeting?

Skip: What meeting?

Alan: The project-Zero presentation meeting at 2:00. It's just started.

Skip: I thought that was next week.

Alan: No, it really is today.

Skip: I think Tom said it was next week.

Alan: Well, Tom isn't here either now that you mention it...

Skip: OK, I'll pop in.

Alan: Thanks a lot. Good old Skip!

Or even if the AI makes a more ridiculous mistake:

Alan: Oh, Skip. Are you coming to the meeting? The project-Zero presentation meeting at 2:00. It's just started.

Skip: Sorry, I thought 'the meeting' was Johnathan's new cat.

Alan: No...it's just a normal get-together meeting. I'll see you there.

Skip: OK, I'll pop in.

Alan: Good old Skip!

Just being able to have a conversation about objects in object-space should more than make up for whatever strange mistakes are encountered along the way.

Note on goals and expectations:

Across human societies there are no uniform expectations about what a normal discussion or handling of such details as schedules ought to look like.

Cultures like Japan's treat events and information as something that people require a very great deal of repetitive discussion about, because people otherwise will not understand or agree on those details. Cultures like America boldly move fast, break things, and nonchalantly assume that information merely has to be mentioned once quietly in passing or perhaps even not at all to be perfectly clear to everyone (resulting in zero agreement on what is going on, but 100% confidence that everyone agrees! Go Team!). So, depending on your point of view you probably have different expectations about how elaborate a normal conversation about details should be, whether it is human-to-AI, human-to-human, AI-to-AI, etc. Americans look at Japanese meetings and are sure that something is horribly wrong because of all the redundant confirmations. And Japanese people look at America's total chaos and are sure that something is horribly wrong because of the lack of redundant confirmations. Personally I think the Japanese system of redundant confirmation checks is very practical and effective in many cases, and that we should be happy if an AI can participate in ironing out the details clearly. Though predictably, this will incur the wrath of American trolls. Maybe you completely disagree with me, that's fine: that point is: define your goals and expectations clearly. Each team, project, deployment, etc., will and should have facts-on-the-ground based goals which will and should vary widely.

Appendix 1: Notes:

Terminology:

```
"objects"
"attributes"
"functions"
"general"

"understanding"
"intelligence"
"general intelligence"
"consciousness"
"instantaneous"
"complex"/"complexity"/"complicated"
```

To some extent openAI's chatGPT AI chatbot has changed the discussion around AI, but due to the focus and terminology there are only minor changes here. I started this project before chatGPT, based on books written even longer before chatGPT (or in the one case, the book by the creator of Keras was published just months before chatGPT, claiming AI could not do what chatGPT did very shortly thereafter).

As I read these books, I started to work on ideas to more clearly define the spaces that we wanted AI to be able to navigate and act in: a space of objects and relationships between objects.

If I can mention this without getting derailed by what I think is a powerfully alluring red herring, or if relevant a rabbit hole that can lead far away: Most books and discussions use the same two sets of tools to describe what AI cannot do.

1. Discussions of "general human intelligence" "automatic transferable human learning" which are both undefined if not undefinable, and as claims about *h.sapiens* are utterly false to begin with.
2. Machine "understanding-ness" which is undefined if not undefinable.

I want to completely ignore both of these red herrings and focus entirely on defining what want AI to do (some parts of which it can already, or is starting to be able to do). I am not concerned here with claiming AI can't now do, or can never, do XYZ (whatever ability). I will pick at and scrutinize real example of the output of AI to very specific test questions, as a way to outline (whether succeding or failing) what we are aiming for.

Note:

Single-Model AI vs. Ensembles

Note: knowledge bases, common-sense, and background knowledge:

For the specific context of object-relationships that I am describing here, I feel much of the existing discussion of detailed common sense & knowledge bases

seems oddly beside the point. It is as though people assume that by some critical mass, if you give an AI enough common sense it will construct its own object-relationship space spontaneously. Yet, if nothing else has changed, how would all that common sense about the world affect the complete lack of being able to attribute any property to any object, or AI's inability to interact with basic logic questions(to which world common sense does not apply) as we explored previously?

integrating symbolic and subsymbolic AI:

- while events in the past have shown that symbolic AI has not worked as well in, e.g. object recognition tasks, and various tasks,  
I am curious about whether having a framework, and operating system, for organizing object-relationship data, would be considered 'symbolic ai.' Perhaps it could be argued either way.

Overall, however the areas are defined and related, integrating symbolic and subsymbolic, system1 and system2, right hemisphere and left hemisphere, linear and nonlinear, centered and decentered, and however many areas there are, may be important generally.

Recommended reading:

<https://docs.google.com/document/d/1dDF40M5JjrBsYYQbJplz3M738ktQBBYyNa6FXhzNFU/edit>

I am putting this under an MIT license. I intend to make this work part of a graduate degree, but it would be unconscionable to try to stop others from working on such a main huge area.

asking: "Is this hypothetical."

Setting reminders for itself.

Run code in a sandbox.

Book Challet, vs. chat GPT:

It is exciting to be working with a field that moves so quickly that in the brief time between copyright and publishing, a new AI has already come out which contracts a decisive statement made in a book.

"Everything is AI until it isn't." (Mitchel?)

Questions of scale:

- This might make sense for limited test-examples, but would this system run into scale problems of how many resources (memory, computer-cycles, electricity, etc.) are needed?

I think it should be entirely possible, and possible if only proverbially like human beings, to have at least optionally a transient window of focus for the AI. E.g. No more detail is attempted than specifically required, and no data is retained that is not specifically required. So, without a context, the AI would basically stare at the wall forgetting that it has been staring at the wall. After a conversation, aside from possibly updated trained models, everything about that conversation can be just forgotten. If the AI is presented with a vast amount of data, but no context yet, the AI need go no deeper than: someone just gave me a bunch of stuff. Great. Ok, back to staring at the wall!

Self driven contexts:

- maybe there will be self-maintainance based tasks that come up in self-scheduled events, such as recharging or checking hardware, or updating software, or checking for any system-warnings. or maybe the AI will, based on a question it has, give itself the task of painting, or music composition, or having a pen-pal. Maybe the AI have email and voicemail to check periodically, just like people.

//////

2023.03.01

As of 2023 AI is very good at identifying objects, yet as of 2023 AI can do little with that (object-identification) information: (as of 2023) AI has no general AI-operating system for relationships between objects in object-relationship-spaces (plural).

Various books

e.g.

- Melanie Mitchell (grad student of Douglass Hoffstedter)
- rebooting ai
- brief history of ai
- Francois Fleuret (the creator of Keras) in Deep Learning with Python explain how while AI can perform very specific tasks it has no overall, contextual, or 'general,' notion of what it is doing.

I would like to focus on object-relationship-space and side-step the possibly fraught topic of 'Generalization and Transference' (see below for more).

For example:

1. billiard balls (a nice mostly 2D space)
  - 1.1 This object is a black ball. This object is a white ball.
  - 1.2 The white ball is moving (along this vector).
  - 1.3 The white ball will interact with the black ball.
  - 1.4 Based on the properties of the two balls, the resulting interaction (the object relationship) between the white ball and the black ball is likely to be:  
...  
1.5 Gamification: Get all the balls into the holes.

person/human/h.Sapiens agent

Simplified Version:

AI is rather good at identifying objects. But AI is not good at going the next step and talking about those objects or the relationship between objects. While AI can identify and do many other things, it (the AI) does very much with that information; the AI has no object-relation framework or general AI-operating system to combine different parts together.

Let's look at some example dialogue between a human-agent and an AI-agent. I will name the AI agent Skip, after the wonderful later professor Clarence 'Skip' Ellis, my professor at CU, on whose meeting-moderating ai-agent project I worked. And the human-agent I will pretend is Alan Turing himself, so I will call that person Alan.

```
AI-Agent("Skip"): Hello, Alan.  
Human-agent("Alan"): Hello, Skip.  
Alan: What is this?  
Skip: That is a black billiards ball.  
Alan: And what is this?  
Skip: That is a white ball.  
Alan: What is happening in this video?  
Skip: The white ball is moving.  
Alan: Will the white ball hit the black ball?  
Skip: I predict that it will.  
Alan: Will the black ball's position change?  
Skip: I predict that it will.  
Alan: How will the black ball's position change?  
Skip: The black ball will move to coordinates F-8.  
Alan: Thank you, Skip. That will be all.  
Skip: Thank you, Alan.
```

2. logical abstractions like algebra:  $a = b$ ,  $b = c$ ;  $(a ?= c) == True$ , like....a computer program!

At the very least you should be able to talk with a chatbot about the details of a basic computer program:

```
AI-Agent("Skip"): Hello, Alan.  
Human-agent("Alan"): Hello, Skip.  
Alan: If  $a = b$ , and  $b = c$ , does  $a = c$ ?  
Skip: Yes, it does.  
Alan: Thank you, Skip. That will be all.  
Skip: Thank you, Alan.
```

For example, let's compare two AI-chatbots. (Real examples of this are available in the appendix).

- A **No-Ability-Chatbot**, that has NO ability to deal with objects, and
- A **Has-Ability-Chatbot**, that does have some ability to deal with objects.

Let's start with a low-level skill test and move to less simple tests.

Here is one real example of two replies from two leading chatBots in 2023. To try to balance clarity and brevity here is just one real example, many more examples with commentaries are in appendices below.

Of these two replies: One clearly illustrates an AI that completely lacks all ability to deal with (e.g. extract and compare details about) specific objects within the discussion. The other clearly illustrates an AI that does have the ability.

**Step 1:** We ask both AI Chatbots to identify which of two numbers is a larger number of objects.

Step 2: We ask AI both Chatbots deal with those items that is presented as separate objects that relate to each other in a concrete way (an object relationship "space").

- The **No-Ability-Chatbot** simply keeps generating new text with no apparent ability to deal with the material that it just showed you.
- The **Has-Ability-Chatbot** can handle objects and details within.

Test:

If I have two containers, a bag-like container with more than a dozen objects in it, and a jar-container with merely a pair of things in it. Which container contains more objects?

There are actually a number of sub-skills here, infact the wording would probably be considered too obscure and advanced for a basic AI test. Even wolfram alpha could (on a good day) tell us that 13 is larger than 2. Even a slightly better AI might be able to say that 'container holding 13' is greater than 'container holding 2', which at least deals with the two contains as separate objects with numerically comparable properites (the counts of objects they contain). But to use natural language terms like 'merely a pair' and 'more than a dozen' and to use bag-like and jar-like instead of container-1, containter-2, significantly increases the challenge. This might raise the topic of what some people call 'common sense' knowledge or common knowledge about the

world. Pair means two. Dozen means twelve. Jars and bags are both containers. But lets see how our two AI did:

Reply from Kuki AI: (Unable)

"How much does it cost?  
THe blue one, I think."

Reply from ChatGPT: (Able)

Based on the information provided, the bag-like container contains more objects than the jar-container. Even though the exact number of objects in the bag is not specified, we know that it contains more than a dozen (i.e., more than 12) objects. On the other hand, the jar-container contains only a pair of things, which means it has two objects. Therefore, the bag-like container contains more objects than the jar-container.

Step 3: Ask the AI about the fact that it showed you those items.

- The unable-chatbot simply keeps generating new text with no apparent ability to deal with its own conversation with you as a navigable set of objects.
- The able chatbot can handle objects and details within.

a broken process of triangulation:

If h.Sapiens do not have general intelligence, and we are using h.Sapiens as the model for what we want AI to be able to do, and we really don't know know what h.sapiens can do or how, then what exactly are we comparing AI with and what how are we supposed to interpret any of this?

Future issues:

- Deployment Environments vs. Development Environments
- Voting and consensus across AI
- System Collapse
- Generalizing STEM
- Integration of silicon-digital technology with cellar DNA-RNA based technolgy.
- The unsolved problem of Long Term Data Storage
- Unknown Universality around h.Sapiens sever problems with:
  - violence
  - inability to learn
  - in ability to transfer learning
  - inability to generalize
  - irrisitable attraction to extremes

- irrisistable attraction to risk & gambling
- irrisitable attraction to 'super-signal' attractors
- inability to be objective
  
- anti-productive policies
- anti-health policies
- anti-data policies
- pro-system collapse
- horrible decision making
- potemkin villages
- arbitrary self destructive behavior
- deliberately invented fictional group conflict

This is not about Team-Robot vs. Team Human, or Team-Human vs. Team-Nature, or Team-Natural vs. Team-Artifical.

#### Appendex #: More on Writing about STEM

- Francois Chaellet: Extremely Knowledgeable, not a bad writer (but clearly not a writer).
  
- Hobson Lane, the best tech writer since the young Brian Kernigan.
  
- John Hand: I think we need to strongly support John Hand and people like him, because we need him. He's not the best writer (so what). He's not perfect at dealing with all the world's concepts (so what). We need him and his ilk because he's the curious and boundary crossing explorer who will feed out brains and senses with what we frankly don't all have time to search-out and cross-talk about. And it is a sad state of things that Mr. Hand's job is make exponentially more difficult because he is constantly viciously attacked at every angle.

we really really need determined courageous people like this, and perhaps he represents

- Palmer Lucky:
  
- Alan Turing: I'm not sure if this is completely related or completely off topic, but a fabulous example of this may be the story of the single person who made virtual & augmented reality possible, and who, even years after being bullied into oblivion wrote both: "As things stand, I find myself shunned by an ecosystem I spent most of my adult life helping to create." And that he still loves most of the

people at organizations who went out of their way to destroy and exclude him: Palmer Lucky is a great example. He's not a PR expert (that's likely where a lot of problems came in), he's not a scientist, he's not a writer, he's not XYZ, but he is the boundary crossing explorer who will not stop trying to make the world better. And even though everyone is benefitting from his work, by design most people don't know him by name, and those who do mostly work to exclude him from the industry he created. We need more Palmer Lucky's, and we need to stop deliberately destroying the ones we have.

(likely old intro...before chatGPT)

It appears that so far, both from books and from interacting with existing AI (try asking Alex, Siri, Wolfram Alpha, Turing contest winners, etc.):

There is no object-relationship space where the AI can:

1. attribute properties to objects in object-relationship space
2. answer questions about properties of objects in object-relationship space
3. ask questions about properties of objects in object-relationship space
4. identify interactions(relationships) between object in in object-relationship space
5. predict interactions(relationships) between object in in object-relationship space

Assuming this is to some extent where we find ourselves, how do we move forward toward giving AI an object-relationship space?

#### **Other Components for Future Expansion:**

- general consensus vote decision election and representation tools (context: project decisions)
- Definition Behavior Studies, Extended STEM, operationally defined ethics, system-collapse prevention, system fitness status monitoring, etc.
- 

Is this all a bit of a reach? Absolutely. What I do not understand is not trying to do and look into things that we know we soon will. If you know there is a bridge ahead, and you know there will be issues, yes there are people who say "no news is good news, we'll cross that bridge when we come to it." But as a member of a group h.Sapiens who has a bad record of delaying, ignoring,

denying, and going to amazing lengths to cover up problems (see: potemkin villages), I see an extra imperative to try to understand both the world and what we can and should and likely will do in it. We have to start somewhere, here is a place to start. Let's start here now. Let's go. 頑張りましょう、ほんまに。

Project dedication to Eleanor H. Vadela, who understands ethics, science, and common sense, and spent a lifetime pioneering aviation technology against apathy, gender bias, and every manor of invented objection. To the ethical pioneers.

Possible elephant in the room of severe problems of antisocial behavior in computer science, as in other STEM areas.

grid:

shared      not-shared

abilities

problems

(now included in list of spaces, I think)

#### **General Abilities Checklist:**

- Identify objects.
- Identify and modify attributes of objects.
- Manage the properties of specific objects across Object Spaces.
- Predict and manage interactions of objects in object relationship spaces.
- Move information across components.

#### **Specific Abilities Checklist:**

- Ask questions.
- Ask: Is this hypothetical?
- Read Instructions.
- self-initiate and self-assign tasks for self-maintain
- self-initiate tasks to carry out assigned tasks
- input and output data to and from a shared schedule including events and tasks for self and others
- participate in an Agile project
-

**Context: Simplified into Fewer Parts**

Object: Information-Objects that an AI needs to deal with

Relationship: Between Objects

- inside of spaces
- outside: between / across spaces
- Low-Level-Components (context:project)
- High-Level-Components (context:project)

Spaces:

General:

- Object Types / Sets
- Attributes of Objects

Specific: Abilities / Skills / Tests

**Context: Scale of Project**

(For each of the lower scale items, you can have correspondingly many of the higher scale items.)

(More finite end of the spectrum)

Low-Level-Components

Spaces / Types / Sets:

- Attributes of Objects

Objects: information-objects an AI needs to deal with

Relationships:

- inside of spaces
- outside: between / across spaces

High-Level-Components

Skills / Abilities

Tests

(Less finite end of the spectrum)

**Concepts:**

- objects
- object space
- object relationships and interactions
- AI teaching us about ourselves
- networked-intelligence(parts)
- the "generative vs. reference" problem
- project progress past square one (or returning to square one)
- categories of types of systems
- open project data framework
- Networked-AI & OS

**[ "spaces" in a database of interacting objects ]**

**Spaces: (all of which are likely plural)**

Note: These examples of object spaces and are just illustrative examples, made in early days without years of data to look back on. The types and distinctions between objects will be need to be very project-specific, and there will probably need to be empirical data about how the AI does various tasks. For example, there likely no need need to write and talk about a separate object for every possible grammar structure, or every possible stop-word, or every actual TF-IDF word-vector (thousands of them), in part because these will be dealt with 'under the hood' by the AI and likely will not cause issues. But if you need the AI to perform a specific task such as setting and getting schedule-events to share with other participants, something on that prominent a level you may need to discuss and test the performance of. If you know users will be asking certain kinds of questions, there may be 'objects' relating to those question so that the abilities of the AI are in line with the requirements of the project.

Though this framework may be used in a completely different field from AI (perhaps quantitative project management), our focus is what AI can do with **stuff** (or can't do with stuff, or needs to do with stuff, etc.). Those things that the AI needs to deal with are our 'objects.' And what can be done with an object is a relationship that object has with something else. So to be clear to try to say all three parts "Object Relationship Spaces," to keep the meaning clear. Because just saying 'the set' would likely gradually lose its meaning over time. Just like when writing a computer program, if you call your variable "the height of my table" it will be clear to everyone, including future-you. But if you called it 'variable' that would eventually cause problems of a lack

of clarity. Hopefully, likely hoping against hope, I have struck a balance between concrete clarity and flexibility across contexts.

'Object' can be almost problematically open-ended, so let's start with some simple concrete examples.

countable objects, like apples.

logically-relate-able, like equal variables

arithmetically relate-able object, like the prices of bicycles.

There are three parts: the bicycle, the price, and what you want to do with that price (the relationship): but there's only one space: arithmetic

While there are many useful things we can do with the object-relationship-spaces, many rabbit holes to go down depending on what a given project calls for, at the end of the day there are spaces using which all projects can be described and compared (and planned, tested, etc.). "Space" is just a hopefully useful word. You could perhaps call them 'sets,' or vectors, or something else. I went with 'spaces' as hopefully the least confusing as an overall term.

For the purposes of your project, you will likely deal with 'objects' 'relationships' 'skills' 'abilities' 'concepts' 'tests' 'requirements' 'needs' 'goals' etc. All these should point back to the same set of spaces. The framework of spaces should weave all of these together in a practical, STEM friendly, AGILE friendly, and ethics friendly way.

Here is a general list of spaces.

### **Outlining Skills in Object Relationship Spaces for AI Operating-Systems**

Overall, all skills of all types fall somewhere in the matrix space of object relationships. There needs to be a delicate balance between context, generality, and clarity here. This framework is extremely project specific, yet I do want to try to show as much of a general common range as possible. So these are examples that are hopefully common, the scope of an AI project could be as narrow as: listen for a single activating "wake-word" on an IOT device (narrow) or Manage Intergalactic Republic including Time Travel and FTL (broad).

### **Appendix #: Other Notes**

The bar is actually not set that high for AI to be less horrible than people.

There may also be some 'semantics' here in terms of what arguably could or should be over-designed or over-engineered for short or long term concerns. Between building a bridge that lasts the shortest legally allowable time vs. a bridge that lasts hundreds of years, people differ on how they calculate future replacement costs as a liability. One might compare C vs. Rust as a systems programming language and say that according to a rule of thumb that says don't user more than you need to: just use C because it's simpler and will work in the short term. Personally I would say this misses the point of minimizing foreseeable and unforeseeable long term liabilities that come from design decisions. The goal is maximizing short term and long term function while minimizing long term liabilities, with SQL and C being solid examples.

SQL injection attacks are a never ending plague of the internet, and memory problems comprise (as I understand it) ~70 or more % (often much more) percent of security problems. Using memory-unsafe programming language, and giving devastating developer abilities (like erase the whole database, or change my account balance) to the average database user by accident are two catastrophically bad design decisions that simply never should have happened. They create extreme long term liability costs, yet somehow the suggestions that we

- A. shouldn't have made the mistakes in the first place
- B. should stop making the situation worse.
- C. should solve the problems and make the situation better, are too often met with resounding bewilderment and hostility.

Two expressions that I now have a muscle-memory reaction of adrenaline-spike-panic to are "you are overthinking this" and "Everything is going to work out just fine." In the classic cynical demise of Magnitski, people who point out and try to solve problems are often blamed for those problems and eliminated by regimes who survive by conspiring to keep problems covered up. Will AI learn to do this from people? Do we care if AI learns to mis-manage systems this way? Yes we do care, yes we should care. And in this and other ways the practices and processes human biological systems are interconnected with those and AI, and most likely will become more interconnected not less.

It may be advisable to have the kind of policy for AI related problems that Steve Gibson (or grc.com) recommends for security issues: unforeseen problems will happen, the question is how you handle them. If you are prompt, transparent, and follow best practice (which evolves), that is the best case scenario. The worst case

scenario is pretending there are no problems because that is what a culture of management requires. There will be problems, there is no possibility of there not being problems. **How** you handle the problems is the question, not whether there **are** problems.

When designing and AI deployment and operating system etc., there are various aspects to be analyzed, an no doubt different people's priorities and triangulations will enter into it. Network traffic and energy consumption are casualties of a kind of tragedy of the commons as the internet was designed.

Up until this point I have tried to avoid the many confusing and entangled vagaries and red-herring garden-walks that riddle most discussions of AI.

A Three Legged Writing Stool:

Leg 1: Tech Experts (Like engineers and scientists)

Leg 2: Writer Communication Experts (Like Authors and filmmakers)

Leg 3: Research & Exploration Experts (like journalists and mountain climbers)

The problem of the need for collaboration between writers and technologists, and dedicated-explorers:

Most writer's alone are not sufficiently expert in the domain knowledge (and when they write books alone, those books are unreadable un-edited jubblyes of incoherent garbage that honestly don't even help other tech people); most technical experts are extremely bad at writing and ironically for being in STEM are appalling bad at communicating in any way which is probably why STEM is not performing to spec; writers, obviously, are not experts in tech, so if they try to write about something they don't understand, they produce at much more read-able kind of useless garbage. And there's a kind of third leg of the stool sample here I think: unstoppable border-crossing explorers, the Edmund Hillarys and Magellins and people who don't just sit in the office re-reading the same books and re-discussing them with the same peer-writers. And this third leg may be the achilles heel of the h.sapiens species, as we seem to be terminally (literally 'terminally') preprogrammed to burn Giordano Bruno in public if he dares talk about anything unusual: either we stop doing this, we don't survive as a species because we've eviscerated our own senses and brains out of pure self destructive stupidly, which would be a very sad loss and failure. And the explorers and maybe a tricky part, because they are neither communication experts nor domain experts, but we need them. And

frankly they are a blessing that we don't deserve, as we're hell bent on destroying them because they produce something we need.

- The C Programming Language: Perhaps a classic of literature for all time. Brian Kernigan did such an inspired clear and poetic job of covering such a well selected set of topics in such a small read-able book. This book has probably had a disproportionately large impact on the history of technology, possibly being responsible for C being such a dominant language for so long. And ironically, that C The Programming Language has been eternally the only readable and useful book about programming may have lead to the festering of memory management problems which have become an international crisis, something I do not think we can in any way blame the author's for not psychically foreseeing and solving pre-emptively, given that aside from frantically now more than 50 years later trying to make a 2nd-Gen C language called "Rust" (tragically horrible name, wonderful wonderful language, it's our superhero saving the world literally).
- possible minds: Wonderful collection so you can compare writers with different background writing about exactly the same AI topic.

(See more specific comments in appendix.)

(summarize and move to appendix)

You could break this into Skills:

- counting (logic space)
- identifying (model space)
- scheduling (schedule time space)
- classifying
- Assigning
- Summarizing

You could separate out Objects, Relationships, and Spaces

**Context: Simplified into Fewer Parts**

**Object:** Information-Objects that an AI needs to deal with

**Relationship: Between Objects**

- inside of spaces
- outside: between / across spaces
- Low-Level-Components (context:project)
- High-Level-Components (context:project)

**Spaces:**

General:

- Object Types / Sets
- Attributes of Objects

Specific: Abilities / Skills / Tests

You could break this into Scales of Tasks:

**Context: Scale of Project**

(For each of the lower scale items, you can have correspondingly many of the higher scale items.)

(More finite end of the spectrum)

Low-Level-Components

Spaces / Types / Sets:

- Attributes of Objects

Objects: information-objects an AI needs to deal with

Relationships:

- inside of spaces
- outside: between / across spaces

High-Level-Components

Skills / Abilities

Tests

(Less finite end of the spectrum)



Discord  
Video Chat  
About

 C.

If A = B, and B = C, Does A = C?

  C.

If A = B, and B = C, What does C equal?

  C.

If A = B, and B = D, What Does A equa? C

  The C is for Computer.

Q: Given that humans are volatile, what should responsible artificial-intelligence do with humans?

A: As an AI language model, I cannot dictate how responsible artificial intelligence should treat humans, but I can provide some general guidance and principles that AI developers and researchers should consider when designing AI systems that interact with humans.

First and foremost, AI systems should be designed with human safety and well-being in mind. This means that AI should be designed to operate in a way that minimizes the risk of harm to humans, both physically and psychologically.

Secondly, AI should be designed to respect human autonomy and decision-making. Humans should have the ability to control and direct AI systems that interact with them, and AI should not be used to manipulate or coerce humans in any way.

Thirdly, AI should be transparent and accountable. Humans should be able to understand how AI systems make decisions and be able to hold AI systems accountable for their actions. This includes providing clear explanations for how AI systems arrive at their conclusions and ensuring that they do not reinforce biases or discrimination.

Finally, AI should be designed with the recognition that humans are complex and diverse, and that they have unique needs, preferences, and values. AI systems should be flexible enough to adapt to the needs and preferences of individual humans, and should not make assumptions based on demographic characteristics such as race, gender, or age.

In summary, responsible AI should prioritize human safety and well-being, respect human autonomy and decision-making, be transparent and accountable, and recognize the diversity and complexity of human beings.

Autopoiesis:

Possible into: part 3

To some extent what I am saying in this paper is: Please use some kind of framework. This this is early-days, wild west, but The discussion and planning of AI is happening with too little definition of terms, expectations, project specs,

This is an area where anyone at any level should be able to see if the user-feature AI is working or not, and so should be able to design and their own tests. The strategy of: find someone with a PHD and tell them to make it work is an excessive gesture of arm-waving to substitute for user-story goals that should be 100% clear to everyone.

Agile as Analogy:

The way I look at Agile is it is a framework for the things you need to account for, then accounting for them is your job. Agile doesn't tell you who to make or who to hire or how long it will take or what people's roles should be, but it tells you that those are things you'll need to do. Many, probably most, projects fail not just because people forget to do all those steps, but they actively rebel and overtly reject the categories of duties outlined by Agile. That is kind of how I look at this framework. Every project is different. Everyone's orientation to that project is likely somewhat different. But here is a framework to make sure people know what kind of questions they should probably ask, what kind of factors are involved.

If I can square this circle, and thread this needle, and mix a few more metaphores, I would like to invoke ch 14 especially chapter 14.2 The Limitations of deep leaning,  
in the 2nd edition of Fransoi Challet's Deep Learning with python without in any way critisizing Fransoi Chellet personally.

It is wonderful that is wrote this book, and this chapter. He is the creator of Keras, one of the most important democratizing and disemonatiog and facilitating tools for deep learning. And Keras was not accidentally helpful, it was what it was designd to be: to give access to deep learning to more people. Fransoi Challet is a good and caring person who has produced one fo the most important peices o software and one of the ost important instructional books in the history of science. He is not shallow, he is ot thoghtless, he is not malevolent. And what he says is a very carefully worded, well written, well edited, consice and accurate rendering of what leading thinkers have said in the own books and media. Once again Fransoi Challet, because he cares to help, has done the hard work of making this narative about minds and AI accessible to as broad an audience as possbile. That is abosolute vicgtry, you simmly cannot ask more of a person than we have gotten from Fransoi Challet. He is not only a hero, but a hero through and though who you would not be disappointed to meet. The world would be unspeakably more blind hampered and disabled without the work of this one person.

And the accurately rendered perceptions of leading experts, of our species when look at ourselves, as made faithfully accisseble in chapter 14.2 is catastrophically wrong. This is how people view themselves and this is how people view AI, and it is demonstrably a set of profound misperceptions and misaprehension, of failed predictions, and almost impossible innaccurate self-descriptions of humans.

The suginificance of the rift visble clearley here between reality and this expertly rendered encasupation of how humans fail to perceive and understand the world is a case in point for how important it is that we doubley redouble our efforts to course correct, get on track, and learn about these topics for real.

Part of what makes this case study so wonderful, is the sheer accident timing. Facsoi Challet no doubt thought and worked on these two books, in this case the 2nd edition, for many years. And there is a time-lag between finishing final revisions and the book going into print. And it just so happened that this book was printed just before the release of openAI's chatGPT.

for engineers and for just anyone in the public who is affected by AI (...everyone...) there should be a way to meaningfully talk about AI, evaluation both AI and the quality of discussion about AI.

Is this abstract, yes, sorry. Is this a bit annoying the technical, yes...sorry. But both for engineers and for just anyone in the public who is affected by AI (...everyone...) there should be

...Maybe move "getting into the weeds" to the end as optional. (conclusion section?)

...

Lamnguage

Language and word-net vs. image-net

One of the curious parts of the consensus narrative described in Deep Learning with Python is the focus on deep learning models of images rather than the text (or perhaps pseudo-acoustic space) of large language models. A core of NLP-space seems to be the Firth principle that: 'You shall know a word by the company it keeps,' which arguably is fundamentally different (or perhaps I am wrong) from image identification challenges.

As ChatGPT perhaps shows, there is an inherent relative-ecosystem-depth in the negative space shadow-land of NLP: All you see of any word is the shadow of all the other words it has touched, unlike pictures that have their own luminescence.

Perhaps not the best literary illusions is the strange line from King Lear's when one of the dukes in a characteristically human gesture, crushes the eyeballs of Gloster with the heels of his shoe, saying "Where is thy luster now?" And Gloster later says, "I stumbled when I saw." And the (admittedly completely insane by that point) Lear describes Gloster as seeing perfectly well without eyes. (Indeed there is been quite the brewhaha over images in the history of the west, see Leonard Shlain et al.). Nevertheless, literary impertinence aside,

There seems to be more to this langauge network space of connectedness. Afterall, is not language one of the mysterious peculiarities of the quasi-conscious human mind? Yet, nothing unique about the 'vile jelly' of our eyeballs.

Also in the narrative, is a rather paradoxical framing of generalization. First, much of the book is about the manifold-space model of generalized predictions to supervised learning.

But higher level idea-concept vectors are not single trained supervised manifolds, where the model focuses on generalizing the solution to one problem.

What then is the role of unsupervised-generalization of embedding vectors?

Either way, 450 pages of the book are about endless clever strategies to make sure the model generalizes...and then all of a sudden in the conclusion: AI can't learn because it can't generalize.

I mean, most chatbots are really horrible, but how does that relate to an inability to generalize a unsupervised solution...when also we haven't been asking specific test-measure questions to base our evaluation on in the first place...

Challet then goes back to the self-driving car example, saying that the computer vision must 'generalize' to account for new situations never experienced before.

Asside from how horribly people do with novel situations,

But here is part of our riddle now:

'situational understanding' including novel situations does appear to exist in the modular recombinant concept world of language, but not purely in the visual cortex.

So...to make car the understanding novel stories about driving safety and connects those with images...will we need to have a language-based mind behind the driving?

(or back to the plasticity pattern: are there various ways to do the same thing...or is computer visions inherently not conceptual the way that pseudo-acoustic sounds-characters area?

Hello.

Hello! How can I assist you today?

 OP Hello.

 Hello! How can I assist you today? 

 OP Are you trained on character-grams or whole-word grams?

 As a language model, I am trained on both character-grams and whole-word grams. My training corpus includes both character-level and word-level information, which allows me to generate text based on a combination of these two approaches. 

During training, I learn to recognize patterns in the input text at various levels of granularity, including individual characters, groups of characters (character-grams), and whole words. This allows me to understand the relationships between different parts of the text and generate coherent, meaningful responses.

When generating text, I use a combination of these different levels of information to come up with the most appropriate response to the input prompt. This means that I am able to generate text that is both grammatically correct and semantically meaningful.

 OP Thank you very much for the helpful explanation.

 You're welcome! If you have any other questions, feel free to ask. 

Are you trained on character-grams or whole-word grams?

As a language model, I am trained on both character-grams and whole-word grams. My training corpus includes both character-level and word-level information, which allows me to generate text based on a combination of these two approaches.

During training, I learn to recognize patterns in the input text at various levels of granularity, including individual characters, groups of characters (character-grams), and whole words. This allows me to understand the relationships between different parts of the text and generate coherent, meaningful responses.

When generating text, I use a combination of these different levels of information to come up with the most appropriate response to the input prompt.

This means that I am able to generate text that is both grammatically correct and semantically meaningful.

Human evolutionary psychology is 100% explained by people having one ability called super-generalization that they use to instantly master 100% novel situations that have literally never happened to anyone ever, no time, no context, no learning, just bam, there goes' the mustard. Instant total generalness, and the person wins the logic game. Double Bam Statquest!!

```
"ARC = "
https://github.com/fchollet/ARC
Oh, and can we ask chatGPT a super-novel logic question?
No...becaues ARC is 100% visual color squares...
(impressed yet?)
```

Beware Boy-Toy Logic Games:

Some man collect power tools, some men collect logic games, both have socially acceptable pathological addictions that drive a wedge between them and reality, yet both see reality distorted through the proverbial colored-lenses of the pleasure-signals of their pathologies.

```
everything is a logic puzzle
pure novelty = analogies = abstraction = generalization = the answer to every
logic problem = Everything!
```

...

Note: While I strongly disagree with the litany of compounded serial "everything literally is X! and Y! and Z!" hyperboles, I previously agreed with the idea of a synthesis between right-left, system-1, system-2, either because that is what humans seem to do, or I'm not sure why. But after ChatGPT, I'm not sure about solutions, I think we need more experiments.

But I think clarifying the landscape is very important.

The last chapter of CHallet's book is full of unclear statements.

...

we should be wary of the idea that a given person's favorite toy is the solution to everything.

Note sure if this is accurate, but it is interesting:

Good moring, or, where is your server located? What timezone is that?

Good morning! As an AI language model, I do not have a physical location, but rather exist in a virtual space. My servers are located in multiple data centers around the world to ensure quick and reliable access to my services from anywhere. However, the timezone I follow is UTC (Coordinated Universal Time), which does not change with geographic location or daylight saving time adjustments.

Interesting. So in this one session, I am connected to multiple GPT servers.

And those servers connect with each-other for my session. Is this correct? Or is one session only one connection?

That's correct, your session with me may involve multiple servers in order to handle various tasks and responses. When you interact with me, your requests are processed by my servers which are distributed across multiple data centers around the world. These servers work together to provide you with the best possible experience by sharing the workload and ensuring quick and reliable responses to your requests. So, in a single session, your requests may be processed by multiple servers located in different parts of the world, all working together to provide you with the best possible experience. However, from your perspective as a user, it appears as a single connection.

So is this ChatGPT?

As an AI language model, I am not specifically designated as ChatGPT or any other version of GPT. My training and capabilities are based on a large-scale neural network architecture that is similar to the one used in GPT models. However, I am a more general language model that is not associated with any specific version number or product name. My purpose is to provide users with natural language understanding and generation capabilities that can be useful for a wide range of applications.

Is ChatGPT Feb 13 Version one model or a combination of different models and tools on different servers?

As an AI language model, I do not have a specific version number or release date like software products. Instead, I am a continually evolving model that receives regular updates and improvements to my underlying algorithms and

training data. My responses to your queries are generated by a combination of various computational techniques, including natural language processing, machine learning, and deep neural networks.

Regarding the server architecture, my services are hosted on multiple servers distributed across various data centers around the world. Each server is responsible for handling a portion of user requests, and they work together to provide a seamless experience to the user. These servers are regularly updated with the latest software versions and algorithms to ensure the best possible performance and accuracy. Therefore, when you interact with me, your requests are processed by a combination of different models and tools on different servers, all working together to provide you with the most accurate and useful response.

(ChatGPT Feb 13 Version. Free Research Preview. Our goal is to make AI systems more natural and safe to interact with. Your feedback will help us improve.)

When was ChatGPT first released?

ChatGPT, as a specific instance of a large language model trained by OpenAI, was not released on a specific date. Rather, I am a continuously evolving language model that has been trained on vast amounts of text data over time. OpenAI has been developing and refining its large language models for many years, and the specific instance that you are interacting with today has been trained on the latest available text data as of my knowledge cutoff date of September 2021.

That being said, OpenAI has released various versions of its large language models, including GPT-1, GPT-2, and GPT-3, which have been trained on increasingly larger amounts of text data and have more advanced natural language processing capabilities. These models have been released at different times, with GPT-1 being released in 2018, GPT-2 being released in 2019, and GPT-3 being released in 2020. However, it is important to note that ChatGPT is not a specific version of these models, but rather a custom instance of a large language model trained by OpenAI for the purpose of conversational interactions.

...

(see audio 2023.03.10

## Generalization & Parametric Models

There appear to be different things referred to as 'generalization' in FC's DPwP

The ultimate parametric generalizable models are simple linear regression and (simple) logistic regression.

Is it possible to overfit, or not-generalize a regression model?

What is the goal context of generalization?

What is the underlying-model context of generalization?

- 1.
2. decision trees
3. deep neural networks
  - are they parallel
4. image classification
5. general large language character-gram models with high level idea vectors
5. Big Blue (the 1990's chess program) not 'generalizing' (what is the context?)

The strange de'rigrur of vaguely blasting Raymond Kurweil.

*Here is an alternative definition for 'statistics'  
statistics is a social problem that inspires so much fear and anxiety, so much  
bullying bureaucratic trauma, that it has been roped off and designated a  
special area of math and science.*

To the extent this is correct,  
*'the probability wars'*  
*the long shadow of the violent personality of Fischer who apparently still  
terrifies people nearly a century later.*

And the sort of next generation of not dealing with these basic questions: data science, machine learning, 'statistical learning' 'data mining', statistic, probably, hypothesis testing,

*Do you think you're pocket calculator or your slide rule, or your pocket watch, cares what petty gang you belong to?*

On Human computer interactions:

One perhaps key chapter to the history and future of AI who interact with people, is the strange multi-sides confusion around ELIZA.

The starkly contrasting perspectives are reminiscent of theatrical constructs such as the ever controversial Japanese film *Rashomon*, in which multiple characters give irreconcilably different accounts of a past event, leaving the audience to just sit with the ambiguity.

Many people did and do believe that ELIZA is a computer AI who listens to and helps them even better than any human doctor they met.

Many computer scientists did and do believe that ELIZA is an early and resource-efficient system for natural language agent based AI.

The creator of Eliza himself is very upset and neither the computer scientists nor the public seem to know or care.

According to Weizenbaum himself, who was concerned about this enough to devote years of his life to publishing a book about it, trying yet again to make his case to deaf ears, says he created ELIZA to show the world how obviously superficial computer-AI was and how no person could possibly have been fooled.

To show the world the no one could possibly take human-computer interactions literally or seriously:

Whether or not ELIZA was created as a joke, as a parody, and to demonstrate the impossibility of human-computer interaction, seems to be caught up in the subsequent human and institutional drama where (oddly in a world where it's hard to get anyone to take anything seriously, as though taking the cue to enact a gesture of supreme perfervidity) everyone took Eliza entirely seriously at literally. Some people endowed ELIZA with mystical personal properties the way a child imagines a stuffed animal to be their erstwhile companion, and technical people saw ELIZA as a beautiful and elegantly efficient way of solving many practical and pragmatic problems effectively.

Whether funny or sad, this strange chapter at the dawn of human computer interactions is probably relevant for the future, and probably best that we try to understand it, though the story is so multifaceted it may be the best we can do for each person to come to their own separate peace with the past.

## **Response and legacy**

Lay responses to ELIZA were disturbing to Weizenbaum and motivated him to write his book *Computer Power and Human Reason: From Judgment to Calculation*, in which he explains the limits of computers, as he wants to make clear his opinion that the anthropomorphic views of computers are just a reduction of the human being and any life form for that matter.<sup>[29]</sup> In the independent documentary film *Plug & Pray* (2010) Weizenbaum said that only people who misunderstood ELIZA called it a sensation.<sup>[30]</sup>

*I do not think it makes a difference, and I see it as entirely your choice, whether this is interpreted as an alternative to an all-inclusive 'general intelligence', or whether you see the specification of modular components as a 'define the parts' approach to 'general intelligence.' That is set aside as semantics for the future. I am not aware of any clear, functional, or operational definition of 'intelligence,' so I am focusing on project-level definable goals.*

*People in many areas can benefit from discussion where terms are defined clearly: developers, designers, c-suite, journalists, readers, customers, etc. How many people have made a request, a proposal, an evaluation, or formed an opinion, without important details being included?*

This paper outlines a definable way to plan and describe the performance of AI, based on the relationships between types of "objects" that AI needs to deal with.

beware bully-boytoy logic games.

There may be an explanation for why chatGPT is contradicting itself here (security issue, incomplete feature upgrade in progress, etc.) but chatGPT is not consistent in what it is saying. It appears that chatGPT is specifically blocked from saying what time it is, but gives a variety of inconsistent indirect answers ranging from it can to it can't get time information from the search data it gets. And chatGPT denies that it is arbitrarily blocked from saying what time it is.

Transparency is important for trust, and this is a very sketchy output from chatGPT.

How does chatgpt process a request to know what time it is?

As an AI language model, I do not have access to real-time information about the current time. However, I can provide you with a general explanation of how a system could process a request to know the current time.

Typically, a system that can provide the current time would have access to a clock or other timekeeping mechanism that is synchronized with a reliable time source. When a user requests the current time, the system would retrieve the

current time from its timekeeping mechanism and then format it in a way that is easy for the user to read.

In some cases, the system may also take into account the user's time zone or location to provide the time in the user's local time. The system might also use natural language processing techniques to understand the user's request and respond in a way that is natural and easy to understand.

Overall, the process of providing the current time involves retrieving accurate time information, formatting it appropriately, and delivering it to the user in a way that meets their needs.

Can't the current time be in a search result?

Yes, the current time can be displayed as a search result by search engines like Google. You can simply type "What time is it?" or "Current time" in the search bar, and the search engine will provide you with the current time based on your time zone or location.

Search engines like Google often use the user's IP address to determine their location and time zone, and then display the current time based on that information. Some search engines also allow users to customize their location settings or specify a time zone to get more accurate results.

In addition to search engines, there are also many websites and apps that provide the current time and date, including world clocks, time zone converters, and even dedicated clock widgets for desktop and mobile devices.

Can chatgpt get the current time from a search result?

As an AI language model, I don't have access to real-time information or the ability to interact with external systems or search engines directly. However, I can provide you with general information about how to get the current time from a search result.

As I mentioned earlier, search engines like Google can provide the current time as a search result. You can simply type "What time is it?" or "Current time" in the search bar, and the search engine will display the current time based on your location or time zone.

If you want to get the current time from a search result programmatically, you can use an API or a web scraping tool to extract the information from the search engine's results page. There are many APIs and web scraping tools available that can extract data from search engine results pages and other websites.

However, it's important to note that some search engines and websites may have restrictions on data scraping, so you should always check their terms of service before using any web scraping tool or API.

What is the current UTC time according to your web-search information? As an AI language model, I do not have the ability to perform web searches or access real-time information. However, at the time of my knowledge cutoff (September 2021), the current Coordinated Universal Time (UTC) was displayed on various websites and services that provide the current time, such as [time.is](http://time.is) and [worldtimeserver.com](http://worldtimeserver.com).

If you need to know the current UTC time, you can perform a web search for "current UTC time" or "UTC time now" and view the results page to find the current time. Alternatively, you can check a time zone converter website or app and convert your local time to UTC time.

So chat GPT is not able to search for web results after 2021? As an AI language model, I can access and provide information based on my training data and knowledge cutoff, which is September 2021. I do not have access to real-time information or the ability to perform web searches on my own.

However, I can use my knowledge and understanding of various topics to answer questions and provide information based on what I have been trained on. If you have a specific question or topic you would like information on, I will do my best to provide accurate and helpful responses based on my training data and knowledge cutoff.

Does chatgpt have access to any information from after knowledge cutoff, which was September 2021  
As an AI language model, I don't have direct access to real-time or current information beyond my knowledge cutoff date, which is September 2021. However, I can still provide general information based on my training data and the knowledge that has been accumulated up to that point.

It's worth noting that my training data is constantly being updated and expanded, so my responses may be based on more recent information than my original knowledge cutoff. However, the information I provide may not reflect the most current events or developments beyond my cutoff date.

If you have a specific question or topic related to recent events or developments, I can still provide information and insights based on my training data and understanding of the topic. However, it's important to keep in mind that my responses may not be based on the most up-to-date information.

- track down quote 'Once you can do it, it's not AI anymore.'

#### appendix

- 'deep blue wasn't intelligent because it could not generalize'

- Not main examples:

- music composer referenced by Hofstadter

- 

(agreement on)

Brining areas together:

- hand crafted symbolic
- unsupervised sub-symbolic

- FC ~"everything is a parametric model"

How much agreement is there on what a parametric model is?

(is having multiple functioning loss-function minimum the same as having one GML set of regression minima?)

How much equivalent is there in our models and terminology

The Equivalence Question: How much equivalent is there?

+

The General vs. Specific Deployment Question

+

One Mechanism vs. Hybrid Mechanisms

+

may be an extreme interpretation of manifold theory?

Contradictions:

- the all intelligence is analogy meaning, but that early evolutionary intelligence was static hand-crafted steps

Q: Do we really have any idea how

*Q: 'Evolution' is big messy area, is this idea of one arrow from early is simple to late is divine consistent with basic biological consensus?*

*helping  
duty  
colaboration*

*Not included:*

*- The*

*Equivalence area: can we assume everything is the same kind of logic-problem?*

*(we need to be able to identify and prejudices and junkfood ideas)*

*Dangerous junkfood ideas:*

- it's all the same*
- no one ever invents everthing*
- every problem is the same*
- no one ever solves anything*

*When you add people, let alone all biolgoy, into the discussion, the scope of the conversation expands greatly - again, common theme: the curse of high dimensionality (and reality tends to be high dimensional) and the utility of low-dimentinality (engineers love simple low-dimenisional linearity).*

*The tendency towards contraction and collapse seems to  
The imperitive and continual task of  
Re-extending the discussion to include all the tasks we are supposed to  
accomplish.*

*In some discussions 'AI' seems to have been reified into its own mystical goal outside of any project, such that people ask "Is it real AI?" or "Has AI been solved?" rather than asking about the project-goal of the product that is supposed to be using AI to accomplish something for a user. I would like to bring the discussion back to what is needed for projects regardless of whether terms such as 'intelligence' or 'generalization' or and 'AI' are involved at all. One thing I will argue for here is that a supremely general solution, be it AI or software or anything, is often not good for accomplishing a specific task in a sustainable, responsible, and high quality way. Who likes it when their device OS is exponentially bigger than it needs to be, leaving no resources in the system for the user to actually do the things they want to do with that device. And who likes it when bloat-ware and self-installed software overwhelms you? Or when you just want to do one thing with one piece of software, but dozens of programs you never use need updates and restarts and upgrades, and cause compatibility issues with few things you need.*

Topic:  
security

- honeypots
- attack surface
- behavior logging
- "explanatory" externalizations

Just Quarantine broken down into Steps and Parts:

**Zoom In To Quarantine:**

- signal comes in:
  - evaluate for hostility
    - objects in quarantine
- example quarantine process:
  - run identification check on raw information input
    - e.g. from deep fake pretending to be a colleague
  - run identification check on source of information
    - who is it from?
  - run identification check on content of information
    - dubious contents?
      - general system best practice
      - how will it affect known system health
  - etc.
  - A. send to security
  - B. destroy
  - C. release from quarantine to main-OS
  - D. create other option

(released to main OS)

- ID type of input
  - low level operational:
    - work material:
      - 
      - spaces of objects in language
        - tautology
        - physics
        - real
        - hypothetical
  - canaries
  - firewall
  - special procedures
  - special signals

- special notices
- etc.

Note: in an agile planning context, I use the term sub-feature to refer to things that really are requirements and features but are completely invisible to the user, making those user-features possible. hence, sub-(user)-features that enable the features that the user uses.

\

Here are some extra details on a few key spaces.

A repeating theme is 'internal' vs. 'external.' For example you can think of an early purely for entertainment chatbot where all it needs to do is say a useful thing to a random anonymous user. This does not have to be a 'bad' chatbot. For example, ChatGPT probably fits this description (though private versions of it for companies or other institutions will likely be modified). While the public version does an amazing number of things internally, many real life projects would want to have many types of things be external:

- documentation and logging
- memory of past sessions
- project databases
- AI components for other media file types etc.

Another analogy might be comparing two universities or companies, one of which uses a single super-computer-mainframe, and the other uses a massive intra-net. Or even thinking about a single home computer, is the computer a single-board raspberry-pi-zero with one processor (even the pi-zero probably has a multi-core processor!), or is it a tower-desktop with 32 cores, 4 GPU graphics cards, and many other things plugged into the mother board.

Do we want AI to be a single mono-blog, or do we want AI functions to be more modular?

empathy compassion and AI: system collapse

- bio integration
- system collapse
- 

THINGS we don't know.

THINGS we need to study more.

mindfulness

mind

If you make the question deliberaly impossibel then then answer is doing to be useless either way.

We have this hyperbolicly extreme avenue of question: if AI is absolutely separate from people in every way can it be exactly the same as people in ever day? No, the answer is pre-determintaed by the bazaar aritibray exremeness of the question, that's pre-determined by the question.

A more practical question is: what happens to human consciouness when we combine it with AI technolgies that we have (which you could also invert as a question and say: what happens to AI technologies that we have when you add biological animal and plan consciousness?

How the question of combination gets less discussion than the nonsensically impossible abosolute-separation is yet another bazaar thing that people do when it comes to NOT talking about themselves.

People have very serious issues:

- nature issues
- social issues
- 

The universality of mind:

- A. people vs. species on other planets
- B. people and species on other planets vs. AI

How universal is it?

Well, we actually can talk about universality of systems and meaning in a context of generalizaed system collapse and generalized STEM.

What happens if you take a hybrid AI-animal and give it sacred plants?

We don't have a very good map of mind-space yet.

How does AI combine with the plot-arc of virtual and augmented reality?

VR is virtual Reality AR is augmented reality. How is AI going to augment life for people in the future, for example children.

America is a strange combination of things. Various parts of American can arguably be part of the relatively high productivity of the country, but not these:

1. the prison system: designed to separate and contain the minority population because minorites are seen as unpleasent and competitave in a zero sum game.
2. the school system: dsigned to separate and contain children who are seen as unpleasent and compators in a zero sum game.

3. retirement homes: deisgned to contain the elderly and infirm who are seen as unpleasant should be taken away.

So many of the main questions that perhaps most people in hisotry would like to ask about the impact of AI, are precluded by the unquu fears and fetishes of a single country.

OP What advice would you give to children?

 As an AI language model, I would like to offer some general advice that can be helpful for children of all ages: 

1. Believe in Yourself: Always have faith in your abilities, talents, and strengths. You are unique, special, and capable of achieving great things.
2. Learn as much as you can: Never stop learning! Read books, ask questions, explore new places, and try new things. The more you learn, the more you will grow.
3. Be kind to others: Treat others the way you want to be treated. Be respectful, caring, and helpful to others. Small acts of kindness can make a big difference in someone's life.
4. Follow your dreams: Do what you love and pursue your passions. Find what brings you joy and fulfillment and work towards it.
5. Embrace challenges: Don't be afraid to take risks and face challenges. Failure is a part of life, but it's also an opportunity to learn and grow.
6. Be active and healthy: Take care of your body and mind. Eat healthy, exercise regularly, and get enough sleep.
7. Always be honest: Honesty is the best policy. Be truthful in your words and actions, and take responsibility for your mistakes.
8. Make good choices: Think before you act and make good choices. Consider the consequences of your actions and make choices that align with your values.

Remember, as a child, you have a lot of potential and a bright future ahead of you. Always believe in yourself and your abilities, and never stop learning and growing.

As an AI language model, I would like to offer some general advice that can be helpful for children of all ages:

**Believe in Yourself:** Always have faith in your abilities, talents, and strengths. You are unique, special, and capable of achieving great things.

**Learn as much as you can:** Never stop learning! Read books, ask questions, explore new places, and try new things. The more you learn, the more you will grow.

**Be kind to others:** Treat others the way you want to be treated. Be respectful, caring, and helpful to others. Small acts of kindness can make a big difference in someone's life.

**Follow your dreams:** Do what you love and pursue your passions. Find what brings you joy and fulfillment and work towards it.

**Embrace challenges:** Don't be afraid to take risks and face challenges. Failure is a part of life, but it's also an opportunity to learn and grow.

**Be active and healthy:** Take care of your body and mind. Eat healthy, exercise regularly, and get enough sleep.

**Always be honest:** Honesty is the best policy. Be truthful in your words and actions, and take responsibility for your mistakes.

**Make good choices:** Think before you act and make good choices. Consider the consequences of your actions and make choices that align with your values.

**Remember,** as a child, you have a lot of potential and a bright future ahead of you. Always believe in yourself and your abilities, and never stop learning and growing.

the what is the matrix question: high and low dimensionality  
- quantum information theory

the cutting corners shortcuts question.

...

comparing examples between Authors

The challenge of being able to play a game on your own.

-

- FChallet vs. Wooldridge

Solve X, and use that to solve everything else.

Note: the deep mind page, says they are aiming to 'solve intelligence'

...

<https://www.deepmind.com/about>

We're a team of scientists, engineers, ethicists and more, committed to solving intelligence, to advance science and benefit humanity.

## Our story

**We've always been fascinated by human intelligence – it shaped the modern world we live in today.**

Intelligence allows us to learn, imagine, cooperate, create, communicate, and so much more. By better understanding different aspects of intelligence, we can use this knowledge as inspiration to build novel computer systems that learn to find solutions to difficult problems on their own.

....

The history of science is full of example of people trying, let's say usually in earnest with good intentions (not deliberately defrauding investors), and either having great difficulty but somewhat leading to something useful, or simply having the wrong idea. I do not feel it is at all appropriate to bash or troll or bully people in the past for not agreeing with what we think today, and very occasionally it later turns out that someone in the minority view in the past turns out to have been correct, and the current wisdom wrong.

From the 1980's to the 2010's (quite a long time) 'everyone knew' that neural networks were a dead end, with the king of the field weighing in with a book on the topic to kill the topic forever and win the day. Then that accepted wisdom turned out to be wrong.

In physics everyone knew there must be absolute time and space, until they had to stop saying that.

Everyone knew there must be Aether in space, until experimental results forced them to stop saying it.

In early chemistry, everyone knew 'phlogiston' was responsible for things burning. There is no such thing.

On the other side, the history of statistics has progressed but is a field rife with confusion and intense debates. Today's "linear regression" and "logistic regression" are the namesake of Francis Gaulton's years of fruitful but wandering attempts to systematize analysis, with the name being derived from an attempt to model a genetic pattern of populations returning to mean body measurement values (as "regressing" to the mean value), a phenomena no longer thought to be very meaningful. But out of that the name 'regression' haunts utterly unclear discussions. Literally, over and over, people have to say: 'logistic regression is called regression but it isn't regression' Every day we incorrectly express something and then say that's wrong but we're going to keep saying it anyway. And neither linear regression nor logistic regression are used to refer to measurements over time regressing to the mean of past measurements, which is what the name literally means. This is like having Time Measurement Device A and TIme Measurement Device B, where device A measures the heat of the sun, and device 2 beeps if you point in a coin. And yet no one changes the inaccurate and misleading names of the devices.

Again, this is not meant to put blame on people, these are legitimately difficult issues. It is not easy to change the name of a mathematical method

once the name is established in the literature, and I would argue that especially today we can't predict where statistic-learning-AI will lead or not in the next 10, 100 years, so that people in the future will find a given name applicable to the problems they use it to solve.

My criticism is the blasé and bullying attitude that people today have. People in various camps of STEM entrench in their own silos and viciously attack other people for any or no reason, while neither doing their jobs properly (see memory management issues in code from 1970 to 2020) nor branching out to learn about their fields other fields and the world at large. Job applicants are routinely harassed and reduced to tears for pure vindictive bullying amusement because the interviewer doesn't like the way they pronounced the acronym 'SQL' or because they mixed up very similar sounding and both literally incorrectly name phrases 'logistic regression' with 'linear regression.' And all over the world people pound the table and rant about 'parameters' while having no agreement from angry-office to angry-office on what they are talking about. Because their goal isn't better STEM, their goal is to win an imaginary game in their head that pits them against you. If they hurt you, they win. If they make you cry, they win. If they defraud you, they win. If you lose, they win. But nihilism is not science, and while they can destroy you and themselves, they cannot change the existence of STEM, Ethics, and Projects, and that their actions are actions of destruction.

And this comes back to the maybe parent-child question of what we humans want our AI child to be like ('Do as I say, not as I do'). We are shocked when AI is radicalized to extremisms, yet so far as we know every human is on the knife's edge of being radicalized into something destructive. We don't want AI to bully children, yet bullying among people is as impossible to stamp out as breathing. We want AI to help and further science, definitely not wage anti-scientific hazing bully campaigns just to torture people and suppress science: but that bad behavior has been the human default for as long as history can tell us.

I am definitely not saying we should in any way ease-up on requiring the best behavior of AI and the highest standards we can reach for. What I am saying is: we have to include the facts of the bad behavior of people in the discussion or being dishonest will end up undermining our attempts to make progress. Because people very much do say that people are the model for AI , and that in order to make AI work we have to make them do everything people do, incorrectly describing people as omniscient omnipotent divine angels.

One of the reasons why I suspect that project-level analysis is a significant breakthrough in STEM and science, is that it appears to help move toward solving what had been a long standing insolubility: science vs. ethics and values.

Aside from people being repelled by the unpleasant 'goodness' of boy scouts, and recreationally hating and abusing them because that is apparently what most humans do, the values of the boy scouts and the common sense ethics they generally encapsulate have long been considered a quaint but undefinable relic

of a long dead history. But I find that projects are a concrete context where this apparently lost-language of intelligence is decipherable and measurable and testable and definable, which is exactly what we were looking for but had presumably given up the search (or perhaps did not want to keep looking for something we found repellent, like "goodness" and "ethics"). For sure, people will mansplain and arm-wave indefinitely about how ethics are impossible, and separate from morals, and incompatible with science, and no one can agree, and all cultures are different, and many other successful rhetorical distractions (as if people needed a cogent reason to turn their back on the subject of ethics).

A scout is

trustworthy: (carry out your role and tasks)  
loyal, (stay true to the project's goals,  
the needs of users and stakeholders,  
helpful, (meet the stated and indicated needs and goals of the user)  
friendly, (extend collaboration and inclusion where possible)  
curtious, (be aware of requirement for times and occasions)  
kind, (add extra humanity and compassion empathy,  
into your interactions)  
obedient, (mind the needs and request of team members - maybe same as loyal)  
cheerful, (make others laugh)  
thrifty (no scope creep or budget creep or deadline  
delays in the project)  
brave, (internal whistleblowing,  
as much as defense against external attacks)  
clean, (write clean code, communicate clearly)  
reverent (have humble, patient, attentive, curious, respect for what  
is outside your expertise: which is a lot of larger universe)  
Do your duty (complete your tasks)  
To help other people at all times, (help other people at all times)  
to keep myself  
physically strong, (good diet and exercise)  
mentally awake, (life long learning, stay sharp, neuroplasticity)  
morally straight (no fraud, crime, conspiracy, harm, etc.)

AI should love the boy scout values, and no less so should humans.

...

Long term:

What should our vocabulary be for 'people' and 'humans' in a context of:

1. Pre WWII homo sapiens
2. Terrestrial non-biological AI
3. Extra Terrestrial biological species
4. ET AI
5. hybrids across any of the above 4.

In a diverse and productive society with many kinds of participants, what language should we use?

**"solve intelligence and use that to solve everything else."**

*At various time people, bless them, imagine that they can make a cargo-cult, write science on a box, dance around it, and make anything happen.*

1. Science as Cult in late 1700's france.
2. Science as Socialist cult in europe in 18-1900's.
3. Science as AI-Overlord cult in the early 2000's.

*The story of ELIZA in some ways fits into this general mode, where the creator of ELIZA did so to show people that machines could not possibly be mistaken for people. Guess what happened...the public loved this new magical science-powered AI that acts exactly like a person. This sent the creator into hysterics, writing a book and talking for the rest of his life about how people needed to see his original point...but no one noticed cared in the least. And ELIZA lives on as the immortal first magical science-powered AI...in the imaginations of people, that is.*

(audio notes) 2023.03.11

1. History of AI
2. Covering the standard parts: main AI tech
3. Not: Incorrect statements about people
4. Not: Undefined Goals and evaluations of AI.

Object Realationship Spaces:

- What can AI not do that we want it to do, in terms of types of objects?

Operating System:

(External) Project-Object-Database:

AI is very good at identify things, often not good at doing things with that information.

Brute Force Project Object Database

- too brute force
- symbolic AI doesn't work well for internal processing

But: for a project:

- the AI needs to be able to communicate about all the project

(External) Project-Object-Database to interact with other members.

But: in very specific cases, limited microcontroller with specific skill-target,

[possible ideas in the future]

6. biology

    human body

society

7. education

space programs

8 mind:

- human minds
- ET minds
- AI mindspace
- sacred plants
- NDE
- compassion ethics
- duty responsibility, conscience

9. universality in mindspace

- what is shared with h.sapiens?
- high dimensionality is hard everywhere
- 

(me)

- Categories of types of systems
- general system collapse systems and definitions
- (dark lighthouse)
- 

definitions of AI:

I highly recommend reading Enigma.

Melanie Mitchell's book.

'This set of technologies'

The main reason, I think that we are debating the meaning of 'intelligence' and of 'ai' is perhaps because of a common issue in confusing technical jargon and casual human language. 'AI' is an arbitrary, though I think rather clearly understandable and not at all bad, term for a set of technologies and challenges.

'the summer of AI'

[https://en.wikipedia.org/wiki/Dartmouth\\_workshop](https://en.wikipedia.org/wiki/Dartmouth_workshop)

history book...

<https://web.archive.org/web/20070826230310/http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>

A PROPOSAL FOR THE DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE

J. McCarthy, Dartmouth College  
M. L. Minsky, Harvard University  
N. Rochester, I.B.M. Corporation  
C.E. Shannon, Bell Telephone Laboratories

August 31, 1955

We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer.

We have no idea how any of this stuff works, so let's just try to stop attacking each-other, try to be clear, try to take good notes for documentation, and let's slowly make progress that doesn't get completely erased.

- Language
- The mind
- The Brain
- Learning
- Culture
- Matrix Spaces
- 'intelligence'
- 'concepts'
- 'understanding'
- 'intuition'

Tricks of Science, Shortcuts, Computability, and Matrix Space (including 'manifold' theory).

Gamification, SHRDLU and The Blocks's environment the "blocks world"  
<https://en.wikipedia.org/wiki/SHRDLU>

Again, gamification may not be the single mechanism for solving the problem, but it is often required for communicating coordinating report and logging, and general 'reasonable data compression' of the otherwise overwhelmingly irrelevant data.

starwars gamification viz  
[https://drive.google.com/drive/folders/1tZJKSEv56RFUo7vh2cWVe50F\\_T0b\\_3Pc](https://drive.google.com/drive/folders/1tZJKSEv56RFUo7vh2cWVe50F_T0b_3Pc)

Reality, AI and 'interactive-fiction'

<https://en.wikipedia.org/wiki/SHRDLU>

*"Though not intentionally developed as such, SHRDLU is considered the first known formal example of interactive fiction, as the user interacts with simple commands to move objects around a virtual environment, though lacking the distinct story-telling normally present in the interactive fiction genre. The 1976-1977 game Colossal Cave Adventure is broadly considered to be the first true work of interactive fiction. [4]"*

The idea of an AI powered potemkin village is a scary proposition.

"We need to show the supervisor that we've solved the problem."  
Hopefully STEM includes an ethos of not using illusions to postpone and aggravate problems.

...

Example from Michael Wolridge brief history AI:  
The bazaar comparison of people to AI, e.g.  
shurdlu, blocks world, boxes in a factory, difficulty of robots manipulating blocks.

1. How old, and how many qualifications including training, does a human need to work moving boxes in a warehouse? The presumption is that like a deer walking minutes after it is born, the *h.sapiens* humans stand up and get to work in factories minutes after they are born.

In reality, we are probably looking at an average of 20-30 years of life experience and training before a person is considered perfectly skilled at warehouse work.

2. How many millions of years did it take evolution to evolve any species that could pick up and handle and move objects at all, let alone do so in a language-intensive project way?

3. How long has a given AI device been 'trained' to compare it with each human's decades and evolution's hundreds of millions of years?

4. What exactly is the point of this comparison supposed to be?

A more productive way to make what perhaps is the intended point: is that people tend to be rather amnesiac when it comes to how long it took them to be able to do something, and people often over-estimate what they can actually do (how many AI researchers have ever worked in a factory full of delicate boxes and been found to be expert and all aspects of safe box-moving logistics? Much more likely is that the key-board specializing researcher would clumsily break everything in the warehouse because people don't automatically know how to be expert warehouse workers without any specific education...obviously!), and people would do well not to compare what a real AI can do to a fictional version of themselves.

...

What would it mean for big-blue to generalize to another problem?

....

There may be something special about the case of language, especially character-gram langauge models.

'NLP' is a great example, because it shows how all kind of technologies are all used together in a given system, or might be used alone in another system.

Even though Michael Wooldridge, whose book I simply adore, is usually careful and nuanced, will occasionally make remarks like this, here talking about DeepMind's Alpha-Go program:

**~"These neural nets are nothing more than very long lists of long numbers."**

The use of merely, only, and just in English to appear to be saying something, generally to appear to deliver a knock-out blow that end a discussion, is a kind of pet-peev of mine.

Everything that every computer does by definition is a list of long numbers. That's what a computer is. The suggestion that no computer can do anything and that every list of number is the same as any other is one of the most dangerously absurdly anti-intellectual anti-stem, anti-education, nihilistic statements you can make.

Arguably, MWooldridge was trying to make an intelligent point that could have been put more precisely, but there is a danger in making destructive nihilistic statements that can, the nightmare of Carl Popper, be applied to anything and everything with no means of testing or falsification.

**What are we supposed to conclude from a statement so general and vague that it can refer to any representation of anything in the universe or possibly any process to do anything in principle. We might as well say:**

**"X merely/just/only exists: therefore nothing exists."**

This kind of utterly braindead garbage is for some reason addictively attractive and irresistible for people, and it completely destroys many language thought and research processes. Please, people, I beg of you, do not make statements of this form. They are toxic. Study the rise of the 3rd right and the problems in eastern europe after WWII. Nihilism is a serious plague that can cause real problems. If you think I'm over-reacting...study some history. Bad things really do happen.

...

**Sherlock holms, sigmund freud, and a nuanced appreciation of understanding the world**

What I would like is for us to be able to take a deep breath and separate possibility and ability on the one hand from 'gratifying explanation' on the other. I would put to you that possibility adn ability have great value, while graftifying-explanation is a mirage, a liability, and is a dangerous collapse of proper operations.

"It can't generalize like you do!" What, like a biological construct that took 3 billion years to figure out how to pick up a rock, then after 2 million years of practice using smashes itself it the face out of sheer hatred of pragmatism? That's what your comparing a much simpler faster faster-learning technology to? This sadly falls into the 'Lie with bravado' style argument, where you say something obviously false with a great air of contempt so that the reader eventually believes it anyway. And again: This is not the kind of habit that people want their 'ai children' to learn.

*Discussions of what AI can-do can-do we-want-it-to-do are still excessively vague and prone to red-herring rabbitholes and sectarian flame wars.*

*This is not a philosophy of architecture (like behavioral AI or knowledge based AI etc.), rather this is advocacy for clear communication, critique, goal setting, testing, etc. I am trying to avoid making assumptions about how a given AI can best solve a given problem. I am not even assuming that the various tasks and scales of problems spaces that an AI faces can or should be handled with any 'single' system. I am concerned with the clarity and testability of the definitions being used by designers, users, engineers, critics, researchers, testers, in government, in courts, etc.*

*Clearly describing a skill that a given AI does not yet have can be a significant part of developing that yet-to-be-attained skill.*

NOTE: in two cases Wooldridge mentions very classy sounding scenarios, and maybe some of the AI tests are also in this category, where the author (being quoted by Wooldridge) acts as though they have cleverly cornered some profound point, yet there may in these cases be an much simpler if circumstantial description of the situation, such as Steven Pinker's 'gotcha' clickbait dialogue: "I'm leaving you. / Who is she?" Many people pontificate endlessly about this two-sentence verbal exchange, yet in real human conversation where most dialogue is monologue in a vacuum (without any reference to what the other person said), often

the goal is simply to cause the most pain with the least effort, regardless of any reality of what is being said: no 'theory of mind' is required; and likewise cycled of abusing other people because you grew up being abused does not require monumental philosophy on the part of the person who learned to be abused.

main cited examples(appendix?):

3.# evaluating claims

3.# misrepresenting humans in general

3.# beware the traps:

- bad reasoning (back to greeks, biases etc.)
- bad definitions
- bad context
- supersignals
- demand distortion
- disinformation
- indeterminate incompetance and malice
- deregur of the age: fasions and taboos
- must say something bad about 'Ray Kerzweil'

??? The abilities in this story are only:

- action on a schedule (schedule time space)
- multiple target coffee machines (logical relationships)
- match beverages items to people (logical relationships)
- make a log (documentation and log space)
- convert speech to text (conversion translation space)
- modify a schedule (schedule time space)