

The LLM Chatbot as a conversation buddy for Scientific Inquiry

Abel Yang, Ryanne Rae Ang Pei, Ong Jun Kiat Justin, Ren Yuhua, Liu Bi Tao
Department of Physics, National University of Singapore

Teaching Innovation with GenAI (Use AI) - HSI2009 What is a Planet

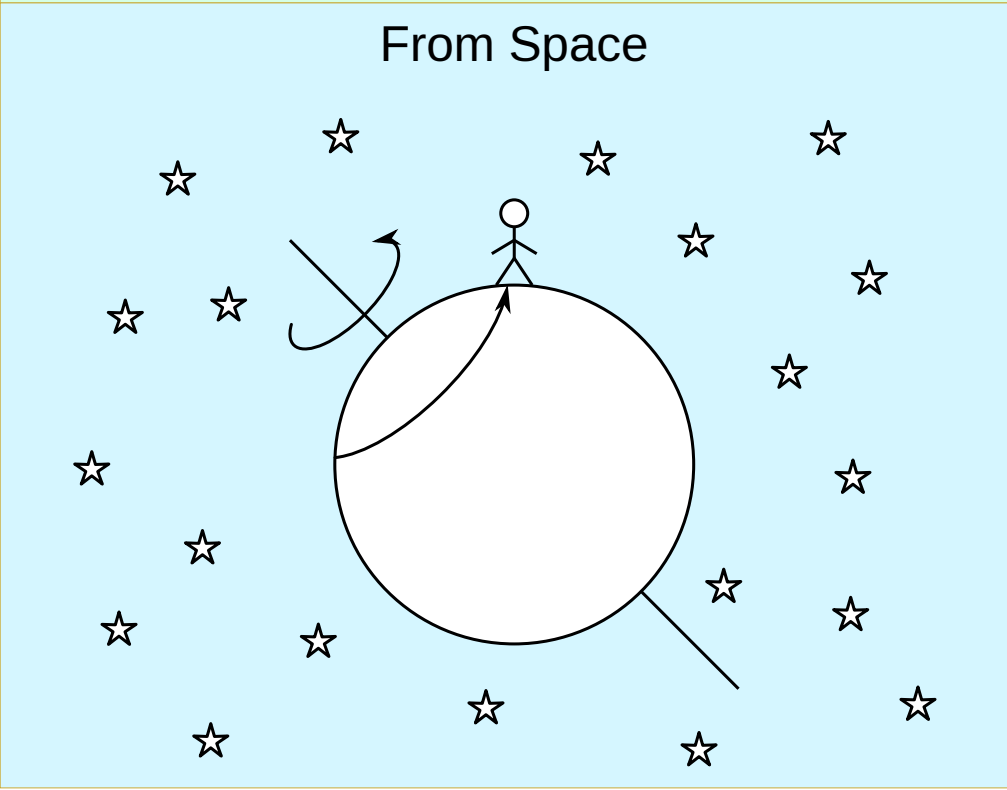
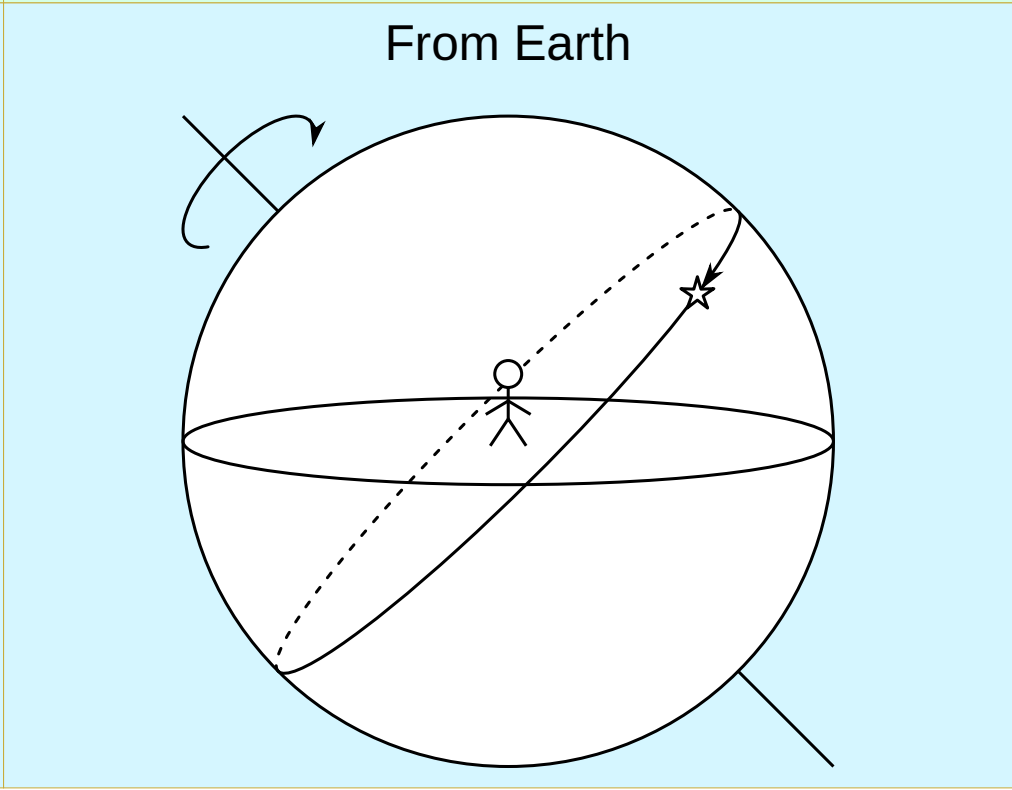
Background

This is an activity set within the course HSI2009 What Is a Planet in the Scientific Inquiry pillar of the CHS Core. Among the goals of the scientific inquiry pillar include "...illuminating the contemporary evolution of ideas in the subject theme, from contested to established science, together with false starts and pseudo science..."

With astronomy as a theme, HSI2009 illustrates the process of scientific inquiry starting from prehistory and ancient efforts to understand the sky. Central to this theme is the two-sphere universe.

The Two-Sphere Universe

The modern accepted description of the solar system and universe is in fact counterintuitive to a naive observer. With no prior background, an observer watching the sun, moon and stars will come to the conclusion that the earth is stationary and the sky moves - This is the origin of the geocentric universe.

The Heliocentric Model	The Two-sphere Universe
The earth is moving (orbiting) and rotating.	The earth and the observer is stationary.
The sun is at the centre of the solar system.	The earth (and the observer) is at the centre of the solar system.
The sun, moon, stars and earth are made of the same material.	The sun, moon and stars are made of different material than the earth because they do not fall to the ground.
	

Motivation

Students were already known to consult (and often rely on) LLM chatbots for a wide variety of topics. However, LLM chatbots are known to make things up especially when the complexity of a session exceeds its context window. As a disembodied body of knowledge, it is prone to misinformation and sometimes fail to pick up context. The chatbot assignment is intended to highlight the shortcomings of GenAI and underscore the importance of empirical evidence and independent critical thinking.

In past semesters, HSI2009 took the approach of in-class open-internet quizzes with the intention of being formative assessments that lead students to think about scientific inquiry in the context of the subject matter. However, being open-internet, many students outsourced their thinking to chatGPT, defeating the purpose of the activity.

The following question was given in an open-internet quiz:
In no more than one paragraph, explain why science should be quantitative where possible.

Many student answers included a paraphrase of the ChatGPT answer of
Science should be quantitative where possible because quantification allows for precision, objectivity, and reproducibility in observations and experiments. By using numerical measurements and mathematical models, scientists can make accurate predictions, identify patterns, and test hypotheses rigorously. Quantitative data also enables comparison across different studies and ensures that results are not subject to personal interpretation, leading to more reliable and universally applicable conclusions. Ultimately, a quantitative approach enhances the ability to detect subtle effects and build a deeper understanding of natural phenomena.

Following this answer, some students also added a misconception that quantitative data is also objective, which is not true: quantitative data can be biased, while the aggregate of qualitative data can be objective.

Chart 1: Quiz question affected by GenAI

With appropriate scaffolding, we can instead use a GenAI chatbot as a debate partner who has read the internet but understands little of it. This activity serves to reinforce the importance of empirical observation, independent verification and the process through which scientific ideas attain credibility. It also serves to challenge students' notion of chatbot capabilities and underscore the importance of fact-checking, building on the skills learned in HSI1000.

The chatbot assignment was first introduced in AY2024/25 Sem 1 as part of a series of assignments. The prior assignments provide the scaffold and context of the chatbot assignment.

The Night Sky

The students were first tasked to observe the night sky with their own eyes, and track the movement of the sky over the course of a few weeks. This includes making at least two sketches of the sky.

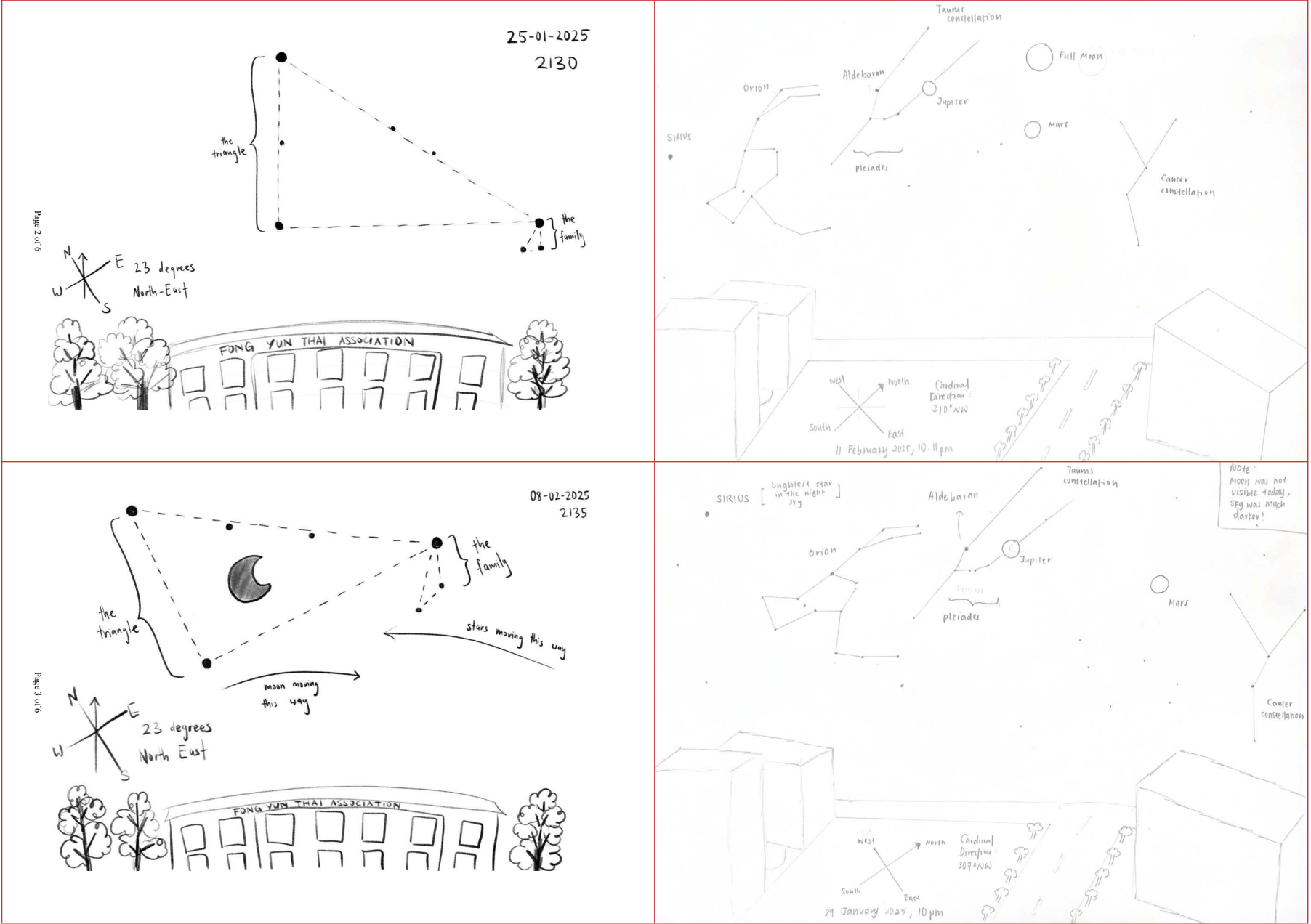


Chart 2: Student sketches of the night sky

Natural Philosophy

In small groups, students share and discuss their observations to establish reality in an empirical manner, and then evaluate ancient theories about the sky. They will conclude that the geocentric two-sphere universe is plausible and reality to a naive observer, and that the modern heliocentric solar system arises as a consequence of modern observations.

Ancient people had various explanations for how the stars, planets and the universe was constructed.

Part 1: Consolidation

In small groups, compare and combine the observations that you made for the night sky activity into a common description of the sky (if you can).

If there are observations within your group that do not agree, find a way to resolve that disagreement (or if all else fails, agree to disagree). Science is a human effort, and human efforts often involve disagreement.

Make notes on your discussion, especially where you disagree

Part 2: Face value

In class, we have covered a number of theories by ancient people:

- the Ionian school
- the Pythagorean school
- the gei lian (盖天) model
- the hun tian (浑天) model
- the xuan ye (宣夜) model

Using the description of the sky that your group has settled on, evaluate each of these theories.

- As far as possible, do NOT bring in whatever you know, but rather work solely on your observational data.
- Given only the information available to ancient skywatchers, would these theories pass the Baloney detection toolkit?
- Are there gaps in knowledge plausible for ancient skywatchers?
- Are there any observations that do not agree with theory? Can you modify the theory to make it work?

Make notes on your discussion, and how your group has evaluated each theory

Chart 3: Tutorial Activity

The Chatbot

Based on their observations, students are tasked to convince a chatbot that the geocentric two-sphere universe is reality - namely that the sun, moon and stars go around the earth, which is stationary.

Then, following the conversation, students would go on to write and submit a 1-page reflection essay on the conversation.

Assignment

AI chatbots learn from the corpus of publicly available texts on the internet, and process that data through a large language model. It might be arguable that a sufficiently advanced AI chatbot is a representation of someone who has read the entire internet.

Thus far this semester, we have established that the Aristotelian two-sphere universe is an intuitive picture of the universe from a naive observer that is confined to the surface of the earth, without modern technology.

Your task in this assignment is to convince an AI chatbot that the Aristotelian two-sphere universe (or any other traditional geocentric two-sphere cosmology) is reality based on your own observations. In the process, critically examine all the arguments that the chatbot makes, using everything you have learned in HSI1000 (or equivalent). Don't simply assume everything a chatbot tells you is correct, they are known to make things up.

Then, write a 1-page (A4, 11pt text, single spaced, 2cm margins) reflection covering the following points:

- Did you manage to successfully convince the chatbot?
- What are the assumptions the chatbot made?
- Did the chatbot make any questionable (or dubious) claims?
- How much of your original assumptions of the universe is based on "received wisdom", as opposed to actual experiences?

In your submission, attach your chat log.

Chart 4: Chatbot Assignment

Reflection and Results

GenAI is used here as a chat buddy and represents someone who has read most of the internet but understands little of it. This was quite apparent in many of the chatlogs where the chatbot was extremely insistent on repeating the generally accepted conclusions of the sky that is common in everyday narrative - in that the earth goes around the sun and not the other way around.

Many students reported that the chatbots simply could not be convinced of the two-sphere universe. The students who managed to convince the chatbot of the two-sphere universe only did so by first establishing the context of a medieval or aristotelian citizen with limited access to information.

While chatbots are known to make things up, students reported that in their experience the chatbot was more prone to simply gloss over the details and instead quote online references. In many cases, the chatbot also managed to confuse two similarly named (but obscure) concepts.

Students could identify the importance of context (that some incorrect conclusions make sense under limited information), and were more cognisant of "received wisdom" where they had in the past taken statements to be true at face value. The overall process of critically examining the arguments from a chatbot while being aware of "received wisdom" did underscore the importance of independent critical thinking (although it is less certain whether or not students will continue to take the effort to keep on doing so).