

Galactic ChitChat

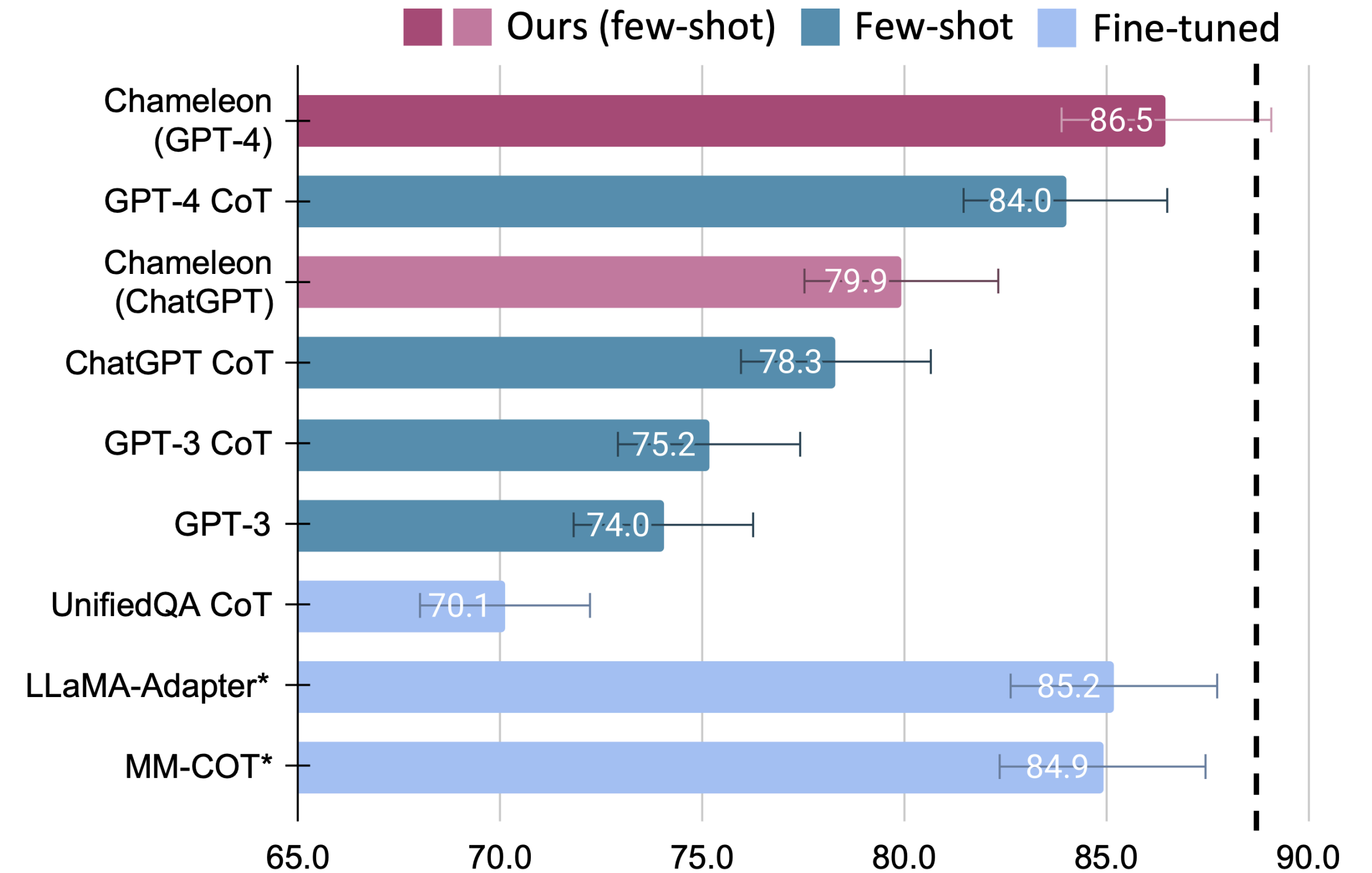
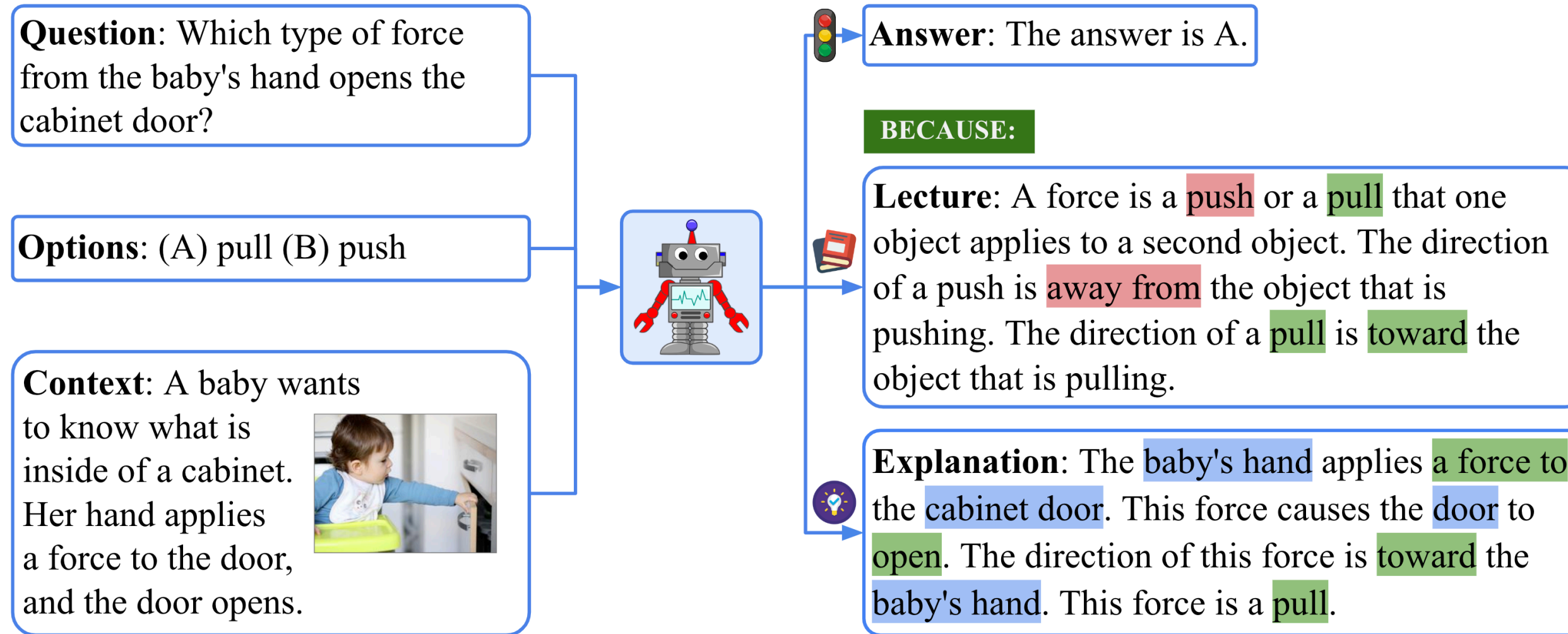
Using GPT-4 to converse with Astro Papers



Ioana Ciucă /w Yuan-Sen Ting (ANU)

KD-IG session, IVOA, InterOp meeting, 9th of May 2023

GPT-4 is SoTA.



Native GPT-4

Summarisation

You: Summarise the main findings from Ciuca et al. 2023, emphasising the quantitative aspects.

I apologize for any confusion, but as an AI language model, my knowledge and training data only extend up until September 2021. I am unable to provide information on papers or studies published beyond that date. If you have any questions about galactic archaeology or related topics based on studies published before September 2021, please feel free to ask, and I would be happy to help! 🚫

Comparative Analysis I

You: What are the main differences between Helmi et al. 2018 and Belokurov et al. 2018?

While both Helmi et al. (2018) and Belokurov et al. (2018) focus on the Gaia-Enceladus-Sausage event, their approaches and data samples are different:

1. Data Sample: Helmi et al. (2018) used a sample of 33,000 stars with high-resolution chemical abundances from the APOGEE survey, while Belokurov et al. (2018) used a sample of 13,000 stars from the GALAH survey with chemical abundances and 6D phase-space information.
2. Analysis Techniques: Helmi et al. (2018) employed a clustering algorithm to identify the Gaia-Enceladus-Sausage event in the 6D phase-space, while Belokurov et al. (2018) relied on the "phase-space spirals" to identify the event.
3. Focus: Helmi et al. (2018) emphasized the discovery and characterization of the Gaia-Enceladus-Sausage event, whereas Belokurov et al. (2018) focused on estimating the mass of the infalling Gaia-Enceladus-Sausage galaxy and understanding its impact on the Milky Way's structure and star formation history. 🚫

Native GPT-4

Summarisation

You: Summarise the main findings from Ciuca et al. 2023, emphasising the quantitative aspects.

I apologize for any confusion, but as an AI language model, my knowledge and training data only extend up until September 2021. I am unable to provide information on papers or studies published beyond that date. If you have any questions about galactic archaeology or related topics based on studies published before September 2021, please feel free to ask, and I would be happy to help! 🚫

Comparative Analysis I

You: What are the main differences between Helmi et al. 2018 and Belokurov et al. 2018?

While both Helmi et al. (2018) and Belokurov et al. (2018) focus on the Gaia-Enceladus-Sausage event, their approaches and data samples are different:

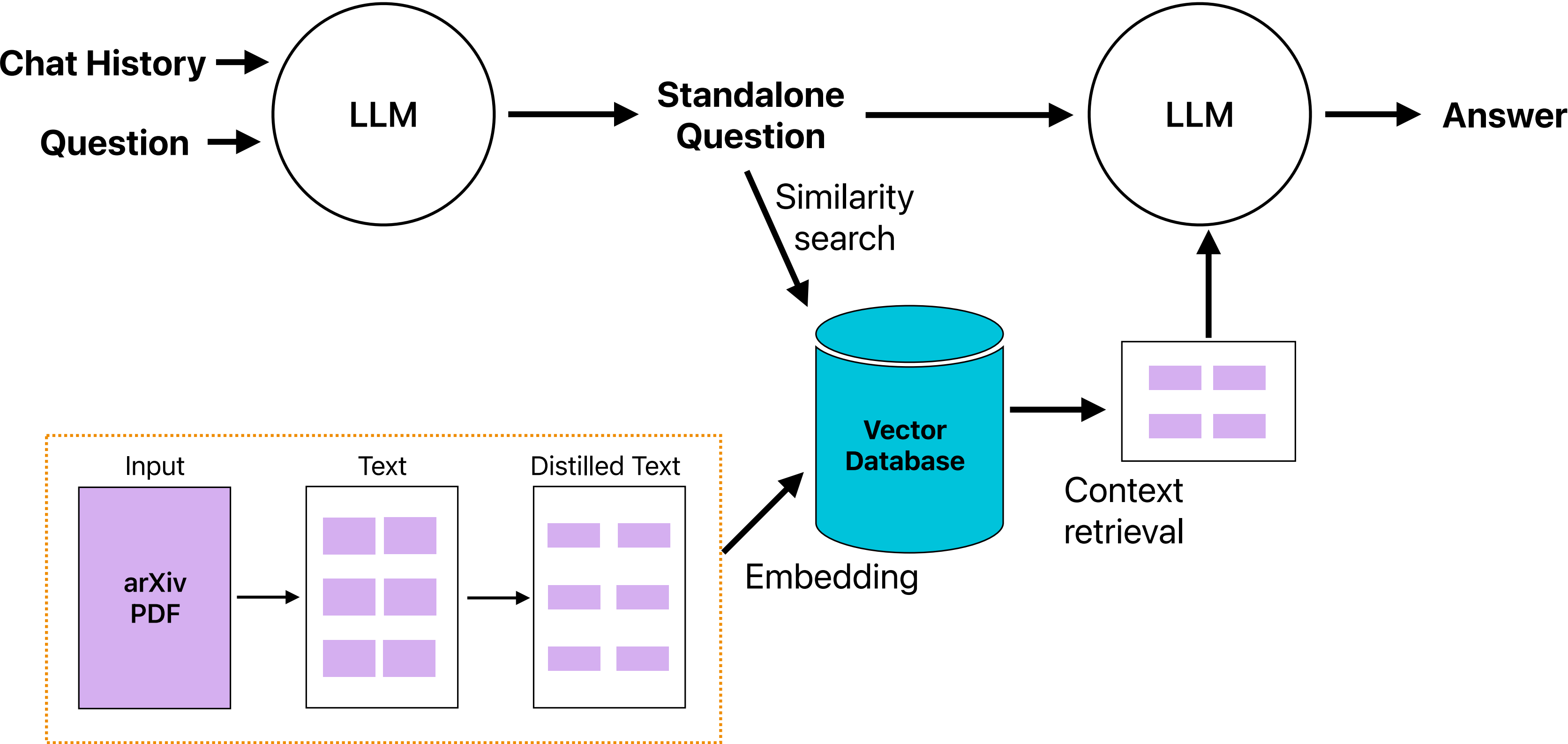
1. Data Sample: Helmi et al. (2018) used a sample of 33,000 stars with high-resolution chemical abundances from the APOGEE survey, while Belokurov et al. (2018) used a sample of 13,000 stars from the GALAH survey with chemical abundances and 6D phase-space information.

2. Analysis Techniques: Helmi et al. (2018) employed a clustering algorithm to identify the Gaia-Enceladus-Sausage event in the 6D phase-space, while Belokurov et al. (2018) relied on the "phase-space spirals" to identify the event.

3. Focus: Helmi et al. (2018) emphasized the discovery and characterization of the Gaia-Enceladus-Sausage event, whereas Belokurov et al. (2018) focused on estimating the mass of the infalling Gaia-Enceladus-Sausage galaxy and understanding its impact on the Milky Way's structure and star formation history. 🚫

Hallucination

Providing context to the GPT-4 model: in-context prompting



Native GPT-4

I apologize for any confusion, but as an AI language model, my knowledge and training data only extend up until September 2021. I am unable to provide information on papers or studies published beyond that date. If you have any questions about galactic archaeology or related topics based on studies published before September 2021, please feel free to ask, and I would be happy to help! 🚫

While both Helmi et al. (2018) and Belokurov et al. (2018) focus on the Gaia-Enceladus-Sausage event, their approaches and data samples are different:

1. Data Sample: Helmi et al. (2018) used a sample of 33,000 stars with high-resolution chemical abundances from the APOGEE survey, while Belokurov et al. (2018) used a sample of 13,000 stars from the GALAH survey with chemical abundances and 6D phase-space information.
2. Analysis Techniques: Helmi et al. (2018) employed a clustering algorithm to identify the Gaia-Enceladus-Sausage event in the 6D phase-space, while Belokurov et al. (2018) relied on the "phase-space spirals" to identify the event.
3. Focus: Helmi et al. (2018) emphasized the discovery and characterization of the Gaia-Enceladus-Sausage event, whereas Belokurov et al. (2018) focused on estimating the mass of the infalling Gaia-Enceladus-Sausage galaxy and understanding its impact on the Milky Way's structure and star formation history. 🚫

In-context GPT-4

Summarisation

You: Summarise the main findings from Ciuca et al. 2023, emphasising the quantitative aspects.

Origins: Ciuca et al. (2023) analyzed the age-metallicity relation (AMR) of stars in different Galactic radial bins. They found an old ($\tau > 13$ Gyr) and low-metallicity population (Babi), a metallicity drop (Dip) around $\tau = 13$ Gyr to $\tau = 12$ Gyr, followed by an increase (Great Galactic Starburst or GGS phase) from $\tau = 12$ Gyr to $\tau = 10$ Gyr. The Dip feature at $\tau = 12$ Gyr exhibits increased [Mg/Fe], while the diagonal GGS displays decreasing [Mg/Fe] from $\tau = 12$ Gyr to $\tau = 10$ Gyr. ✅

Comparative Analysis I

You: What are the main differences between Helmi et al. 2018 and Belokurov et al. 2018?

Origins: Helmi et al. (2018) focus on the merger between the Milky Way and Gaia-Enceladus, examining the age range of Gaia-Enceladus stars and investigating the distribution of its debris in the Galaxy. They provide evidence that the merger occurred around 10 Gyr ago, when a pre-existing disk was present. Belokurov et al. (2018) study the local stellar halo's kinematic properties using Main Sequence stars from Gaia and SDSS. They find that the halo's velocity ellipsoid evolves strongly with metallicity and suggest that a major accretion event involving a satellite with $M_{\text{vir}} > 10^{10} M_{\odot}$ occurred 8-11 Gyr ago during Galactic disc formation, causing radial halo anisotropy. While both studies analyze the Milky Way's formation history and involve Gaia-Enceladus, they focus on different aspects: Helmi et al. (2018) on the merger and debris distribution, and Belokurov et al. (2018) on the kinematic properties of the local stellar halo. ✅

In-context prompting helps with answer quality.

Very early days, but the potential for idea generation is there.

Q: Human: Can you come up with a novel idea that explores a gap in the knowledge provided in the papers you have access to?

A: I propose to investigate the possibility of using machine learning algorithms to identify and classify the origins of globular clusters with higher accuracy, based on their chemistry, kinematics, and age. By training the algorithms on the available data from studies such as Massari et al. (2019) and Helmi et al. (2018), we could potentially uncover hidden patterns and correlations that would help us better understand the assembly history of the Milky Way and the role of massive mergers in shaping its structure. This approach could also aid in identifying previously unknown progenitors and their contributions to the formation of the Galaxy's various components.

Using Agents to interact with our papers

```
multi_tool_executor.run("What are the main findings from Ciuca et al 2023?")
```

✓ 22.8s

Python

> Entering new AgentExecutor chain...

Thought: I need to search for the main findings from Ciuca et al 2023.

Action: Search

Action Input: Main findings from Ciuca et al 2023

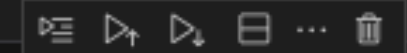
Observation: Our results suggest that the GSE significantly impacted the formation of the Galactic disc, and it catalysed the transition from the thick disc ...

I now know the final answer.

Final Answer: The main findings from Ciuca et al 2023 suggest that the Gaia-Enceladus-Sausage (GSE) event significantly impacted the formation of the Galactic disc. This event catalyzed the transition from the thick disc to the thin disc in the Milky Way galaxy.

> Finished chain.

'The main findings from Ciuca et al 2023 suggest that the Gaia-Enceladus-Sausage (GSE) event significantly impacted the formation of the Galactic disc. This event catalyzed the transition from the thick disc to the thin disc in the Milky Way galaxy.'



```
multi_tool_executor.run("What method did they use? Provide a clear explanation.")
```

✓ 41.3s

Python

> Entering new AgentExecutor chain...

Thought: I need to find the specific method used by Ciuca et al 2023 in their research.

Action: Search

Action Input: Method used by Ciuca et al 2023 in their research on Gaia-Enceladus-Sausage event

Observation: Our findings suggest that the last massive merger of our Galaxy, the Gaia-Sausage-Enceladus, was likely a significant gas-rich merger and ...

I found information on the method used by Ciuca et al 2023 in their research.

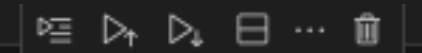
Action: Knowledge Base

Action Input: What method did Ciuca et al 2023 use in their research on the Gaia-Enceladus-Sausage event?

Observation: Ciuca et al 2023 used the BINGO framework to obtain stellar age estimates for 68,360 red giant and red clump stars in the APOGEE-2 survey.

I now know the final answer.

Final Answer: Ciuca et al 2023 used the BINGO framework to obtain stellar age estimates for 68,360 red giant and red clump stars in the APOGEE-2 survey. This method allowed them to study the impact of the Gaia-Enceladus-Sausage event on the formation of the Galact.



```
multi_tool_executor.run("Please describe how Ciuca et al 2023 selected the data.")
```

✓ 52.3s

Python

> Entering new AgentExecutor chain...

Question: Please describe how Ciuca et al 2023 selected the data.

Thought: I need to check my knowledge base for the correct information.

Action: Knowledge Base

Action Input: How did Ciuca et al 2023 select the data for their study?

Observation: Ciuca et al. (2023) selected high-quality red clump (RC) stars with mass > 1.8 M_⊙ and red giant branch (RGB) stars with signal-to-noise (SNR) > 100 from APOGEE-2 data. They then used a Bayesian Neural Network to map stellar parameters to asteroseismic

I now know the final answer.

Final Answer: Ciuca et al. (2023) selected their data by first choosing high-quality red clump (RC) stars with mass > 1.8 M_⊙ and red giant branch (RGB) stars with signal-to-noise (SNR) > 100 from APOGEE-2 data. They then used a Bayesian Neural Network to map stell.

> Finished chain.

'Ciuca et al. (2023) selected their data by first choosing high-quality red clump (RC) stars with mass > 1.8 M_⊙ and red giant branch (RGB) stars with signal-to-noise (SNR) > 100 from APOGEE-2 data. They then used a Bayesian Neural Network to map stellar parameters

- What do we truly want from these LLMs?
- What can we get?
- How can we best evaluate their outputs?