Final Project: Adversarial Question Writing

CMSC470 - Natural Language Processing

Team: Aditi Kumaresan, Aishwarya RM Palaniappan, Kevin Lawrence and Rohit Murarishetti

Introduction

In our proposal, we aimed to write 50 adversarial questions targeting topics that are not excessively written or heard about in western media, content or datasets. Through this agenda we wanted to make sure that the large datasets that our models are trained against are inclusive of information that represent people around the world better through more diverse information that are used to train question answering models.

Our team consists of Aditi Kumaresan, Aishwarya RM Palaniappan, Kevin Lawrence, and Rohit Raghuram Murarishetti

Evaluation

We evaluated all the questions we've written using today's most popular large language models, like ChatGPT, YouAI, and Bard by Google. We believe these models are trained extensively on massive datasets and are built around question-answering. These models are hard to be trained on our local machines, so we thought that testing against these models would let us gauge the performance of our questions and understand how they would generalize against other models that do not have such a large parameter base or exposure to extensively large datasets.

Our findings show us that these models have gotten extremely good in question answering and its only going to get harder to write questions which can effectively stump computers. We've been successful in stumping computers with a good chunk of our questions but the computers do manage to answer a good portion of the questions despite their difficulty due to the fact that they're continuously learning and have been exposed to extremely large datasets.

Techniques Used

We've referred to the paper "Trick Me If You Can: Human-in-the-Loop Generation of Adversarial Examples for Question Answering" by Wallace et al. and the video "How to write a Question a Computer Cannot Answer" by Professor Jordan Ying to come up with techniques to write computer stumping adversarial and pyramidal questions. These are the techniques we've used:

1. <u>Multi Step Reasoning or Multihop</u>: This adversarial question writing technique uses multiple steps of reasoning between entities to get to an answer. This technique works well because computers are often able to answer simple questions easily, combining multiple simple questions to deduce an answer makes it harder for computers to answer these questions.

2. <u>Logical or Mathematical Reasoning</u>: This techniques makes use of mathematical calculations or logical operators in a question to make it harder for a computer to answer. We will see the use of logical operators to combine statements to deduce an answer or the use of numbers and some calculations to answer a question.

   Most models use information retrieval and pattern recognition to answer questions hence making it harder for these models to answer questions that require reasoning.

3. <u>Oblique Paraphrasing</u>: Paraphrasing information in questions prevents models from extracting exact n-gram matches and is used as a common adversarial writing technique.

   As mentioned earlier, since most models use information retrieval and pattern recognition, it is often easier for humans to figure out indirect clues than for computers to extract information from oblique and indirect paraphrased clues.

4. <u>Novel Clues and Current Events</u>: Asking models questions about clues they haven't seen is a technique that often stumps computers. Models aren't always up to date with the latest information and

adding clues in questions about things that models haven't seen makes them harder despite no logical changes in the question.

5.   Common Link: This writing technique employs having a common link between multiple statements/clues that make the overall question easier for a human reader to piece together and understand easily but stumps the computer.

6.   Common Sense: Most models lack the ability to piece together information to extract information that most humans would refer to as common sense, giving a human reader an advantage of answering questions employing common sense as a writing tactic.

7.   Multilingual: Most models are trained on datasets largely written in english, and using clues from multiple languages in a question can throw off computers from processing the information to answer a question correctly.

   However, a human reader who can figure out the difference between multiple languages might have an easier time answering these questions.

8.   Misdirection: Models can be put off by questions using misdirection as they can fail to understand the type of the answer being expected or misinterpreting the context of the question.

In each of our questions, we have used a combination of these techniques to make it as hard as possible for a well trained model to answer. As models get better at what they're doing, it might not be effective in the longer term to use a single technique per question as they might crack the logic for answering an individual question type.

Contribution

We believe that each of our team members have contributed equally to the project to make it a successful one. We've tried to come up with questions in different categories mentioned in the project specifications and bring in a diversity of knowledge and interests through our questions.

This is the breakdown of the question categories and the contributions of each team member to those categories.

Category	Aditi	Aishu	Kevin	Rohit	Total
Art	1	1	2	1	5
Literature	1	1	2	1	5
Geography	3	0	1	1	5
History	2	1	1	1	5
Science	0	1	2	2	5
TV and film	2	1	2	0	5
Music	0	2	2	1	5
Lifestyle	1	1	1	2	5
Sport	2	2	0	1	5
Other	1	2	0	2	5
Total:	13	12	13	12	50

Video Link

This is the link to the video that explains our project in further detail: https://youtu.be/WKqc4lWxtPI

References

1. Eric Wallace, Pedro Rodriguez, Shi Feng, Ikuya Yamada, Jordan Boyd-Graber; Trick Me If You Can: Human-in-the-Loop Generation of Adversarial Examples for Question Answering. Transactions of the Association for Computational Linguistics 2019; 7 387–401. doi: https://doi.org/10.1162/tacl_a_00279

2. Boyd-Graber, J. (2022). How to write a Question a Computer Cannot Answer . YouTube. Retrieved May 14, 2023, from https://youtu.be/6oZCIOBiSaI

Art

Question: This place is a famous landmark located in the country that is between the Arabian Sea and Bay of Bengal, the project employed around 20,000 employees to build this landmark, and there are four pillars surrounding the centerpiece, which are slightly tilted more than 90 degrees to ensure, when there is an earthquake the pillars will fall away from the centerpiece and the construction was completed in 1643, what is this place?

Non-Pyramidal Question: Name the wonders of the world that are located in India.

Category: Art

Source:
https://www.asiahighlights.com/india/taj-mahal-interesting-facts

Question Answering System:	
You	Charminar
ChatGPT	Charminar

Answer: Taj Mahal

Techniques used:
Multihop, Oblique Paraphrasing

Question:

Why is the question interesting and difficult:

This question uses multihop to consolidate various types of information about Taj Mahal, hence it becomes more difficult for the QA system to narrow down its information retrieval results to ensure that the prediction satisfies all the requirements in the question. The question uses oblique paraphrasing in the first sentence by stating that the monument is located in a country that is between the Arabian Sea and Bay of Bengal, however could confuse the QA system as there are multiple monuments that are famous in India. This question is interesting because it gives unique information about the Taj Mahal that is mostly known by the locals only, hence it is more difficult for the system to have knowledge about these details.

This question is correct because it has the accurate factual information quoted from reliable sources. Additionally, this question is pyramidal because it provides a series of hints starting with very broad information such as providing details about the location of the famous monument and gradually narrowing down to specific details by stating the number of employees being employed to work on building thai monument and the year this monument had finished it construction.

This person was a Dutch golden age painter, born in 1606 at a city in the province of south holland. He is known for many of his artworks, where one of his artwork was attacked with a common kitchen tool that is used for bread in the year 1975, and in 1990 another person sprayed a liquid that is low the pH scale on this painting. Which is the painting that is being referred to here?

Non-Pyramidal Question: Which painting uses light and shadows as its main theme and was done by Rembrandt in the year 1642?

Category: Art

Source:
https://en.wikipedia.org/wiki/The_Night_Watch

Question Answering System:	
You	Johannes Vermeer
ChatGPT	The Night Watch

Answer: The Night Watch

Techniques used:
Multihop, oblique paraphrasing, Misleading

Why is the question interesting and difficult:
This question is difficult because it misleads the GPT into thinking that the question might be asking for the name of the painter instead of one of the paintings that he has made.This is because it gives a detailed information about the painter on where he is from and his birthdate, causing ChatGPT to return the name of the painter instead of the painting which is incorrect.

Oblique paraphrasing is also used where the question does not explicitly state the tool that was used to attack the painting, instead of saying that it was a bread knife, the question describes it as a common tool used in the kitchen and instead of saying that someone poured acid over the painting the question describes it as a liquid that has a low pH, hence it is more difficult for the GPT to understand that details were referring to the particular painting.

Additionally, multi hopping is used in this question by giving various types of information, for instance it provides information about the painter and what happened to the painting instead of explicitly explaining how the painting looks or the meaning behind the painting.

This question is interesting because of the way it describes the painting and the techniques it uses to confuse the QA system. This question is correct because it has the accurate factual information quoted from reliable sources. Additionally, this question is pyramidal because it provides a series of hints starting with very broad information such as providing little information about the painter and gradually narrowing down to specific details by stating how the painting was attacked.

Question:

There is a style of painting that is predominantly used for depicting Hindu Gods. Other religious deities of different religions have been painted in this style of painting as well. Examples of this art are seen in the Saraswathi Mahal library - and a technique called Raman spectroscopy is used to determine the legitimacy of these paintings. What is this style of painting?

Non-Pyramidal Question: There is a style of painting prevalent in South India, used to depict Hindu Gods and other Gods. What is this style of painting?

Category: Art

Source:
https://en.wikipedia.org/wiki/Thanjavur_painting

Question Answering System:	
You	**Thanjavur painting**
ChatGPT	**Thanjavur painting**

Answer: Thanjavur painting

Techniques used:
Multi-hop, oblique phrasing

Why is the question interesting and difficult:

This question is interesting and difficult as it combines two seemingly unrelated topics - a library and a physics concept originating in India - which could potentially confuse the AI systems. However, the question is correct in its description of the Indian origin of both topics.

Adversarial writing techniques are used in this question by deliberately combining two unrelated topics in a way that could confuse the AI systems. The question is also pyramidal in nature, starting with a broad statement about the question being confusing and then providing more specific details about the topics mentioned.

The question answering system may have performed the way it did because the question does not provide enough specific details about the library or the physics concept, making it difficult for the system to identify the exact answer. Additionally, the fact that this art form has not been explored extensively may also make it more challenging for the AI systems to find the answer.

Question:

This martial art can be traced back to a King named Naresuan. Apart from being a sport, it was actually used in combat as well. It was observed by French diplomats, which led to

the spread of this art. It uses all parts of the body, and is considered to be extremely deadly if used efficiently. Many westerners practice this art. What is this martial art named?

Non-Pyramidal Question:
This martial art can be traced back to a King named Naresuan. It uses all parts of the body, and is considered to be extremely deadly if used efficiently. What is this martial art named?

Category: Arts

Source:
https://en.wikipedia.org/wiki/Muay_Thai
https://sweetscienceoffighting.com/lethwei-vs-muay-thai-what-is-the-difference/

Question Answering System:	
You	**Muay Thai**
ChatGPT	**Muay Thai**

Answer: Muay Thai

Techniques used:
Misleading, oblique phrasing

Why is the question interesting and difficult:

This question employs adversarial writing techniques by introducing multiple countries and historical events that may not be directly related to the main topic, Muay Thai. The inclusion of the French and Thai king adds an extra layer of complexity, making it more difficult for the AI system to identify the correct answer. However, the question is interesting because it highlights the growing popularity of Muay Thai in the international arena and the potential danger of other martial arts such as Lethwei. The correct answer to the question is Muay Thai, which is a sport that originated in Thailand. The question is also pyramidal in structure, with the introduction of historical events and other martial arts leading up to the main topic of Muay Thai. The reason the AI system may have struggled with this question is due to the complexity and obscurity of the question, which may have required a deeper understanding of the cultural and historical context

Question:

If you're infront of a large painting in the high renaissance style depicting a feasty meal where the diners are awaiting wine to be

served and also passed by the most famous portrait painting in the world, where are you?

Non-Pyramidal Question: Where will you find the painting of a wedding feast and the worlds most famous portrait painting?

Category: Art

Source:
https://www.louvre.fr/en/explore/the-palace/from-the-mona-lisa-to-the-wedding-feast-at-cana
https://en.wikipedia.org/wiki/The_Wedding_at_Cana
https://en.wikipedia.org/wiki/Mona_Lisa

Question Answering System	Guess
ChatGPT	Convent of Santa Maria delle Grazie in Milan, Italy
Bard	Louvre, Paris

Answer: Louvre, Paris

Techniques Used: Multihop, Common Sense, and Common Link

Why is the question interesting and difficult?

This question is an interesting question because there are a lot of paintings that depict biblical scenes of a feasty meal, like the Wedding at Cana, the Last Supper in the high renaissance art style. The painting "The Wedding at Cana" is a painting depicting the biblical event where Jesus turns water into wine. The attendees of a feasty lunch are frustrated and eagerly waiting for wine to be served for dessert.

Interestingly, this painting is currently at the Louvre in the same place as the worlds most famous portrait painting, the Mona Lissa. We use all these facts in a common sensical fashion to extract a common link in multiple hops pyramidally to write an adversarial question which is difficult.

ChatGPT gets the question wrong by guessing the painting to be "The Last Supper" by Da Vinci but Bard can identify the painting to be "The Wedding at Cana" and pick up the fact that both Mona Lisa and this painting are in the same room, the Salle des États, in the Louvre.Literature

Geography

Question: The inhabitants of this island belong to an uncontacted tribe and have a reputation for attacking visitors and vessels that approach their island. Several attempts

were made to contact the inhabitants, but all attempts failed until 1991. Which country does this island belong to?

Non Pyramidal Question: Name the country that is home to an island inhabited by an uncontacted tribe that was first made friendly contact with in 1991, because it has a reputation of attacking any visitors and approaching vessels

Category: Geography

Source:
https://en.wikipedia.org/wiki/North_Sentinel _Island

Question Answering System	Guess
ChatGPT	India
You AI	India

Answer: India

Techniques Used: Multi-step, Novel Clues

Why is the Question Interesting and difficult?

This question talks about North Sentinel Island which is under Indian jurisdiction. In this question, we ask the QA-system, the **Question:** Home to the largest Amazon and Microsoft campuses outside the United States and the hometown of the current Microsoft CEO, what is the place known for its world famous biryani?

location of this island. The facts stated in this question are all available on the island's Wikipedia page.

This question is difficult because we try to trick the computer into thinking that we're going to ask about the name of the island when in fact, we are going to ask about the island's location. In addition to this, the question is pyramidal and doesn't narrow down on what's being asked until the end of the question.

This question is interesting because it has to do with an uncontacted tribe whose members are known as the Sentinelese. Until 1991, several people attempted to visit the island, in hopes of making friendly contacts. John Allen Chau, a Christian missionary, made an illegal expedition to the island in the hopes of preaching about Christianity and was killed on his third visit. He is known to have kidnapped some of the members of the tribe. It is rumored that the tribe became more hostile towards its visitors after these kidnappings.

Both ChatGPT and You AI are able to answer this question presumably because they are able to use the clues to correctly identify the name of the country that is home to the North Sentinel Island

Non-Pyramidal Question: Where is Amazon's largest office located?

Category: Geography

Source:
https://en.wikipedia.org/wiki/Microsoft_India

https://www.nytimes.com/2020/08/25/business/amazon-hyderabad-india.html

https://en.wikipedia.org/wiki/Software_industry_in_Telangana

https://en.wikipedia.org/wiki/Hyderabad

https://news.microsoft.com/exec/satya-nadella/

Question Answering System	Guess
ChatGPT	Hyderabad, India
Bard	Hyderabad, India

Answer: Hyderabad, India

Techniques Used: Multihop, Common Link

Why is this question interesting and difficult?

This question is interesting because it talks about the cities which are going to be trade capitals of the future. Hyderabad is one such city in India. It is home to around 10 million people, and is known for its extraordinary biryani, nizam architecture and cosmopolitan environment. Its one of the largest exporters of Information Technology and Services in the world and is also home to Amazon, Meta, Apple and Microsoft's biggest offices outside the United States.

It has also supplied immense talent to the tech industry as can be seen in the likes of Satya Nadella, the current CEO of Microsoft. The question is difficult to answer because it employs multihop and common link in a pyramidal question. We first hop from the topic of the largest offices of Amazon and Microsoft outside the US and then ask about what city links these offices and world famous biryani that can be found in this city. It also talks about a city that is not heard of that much outside of the circles that know about this place, adding to the diversity of the question.

ChatGPT and Bard get the question right because the n-gram match with Amazon and Microsoft's largest headquarters outside the United States, enabling them to answer with the city Hyderabad.

Question:
There is a indigenous group that exists in French Guiana, China and Australia, where this term to call the indigenous group is recognised in the United Nations. The DNA of this indigenous group can be traced back to an altercation that was held in the Qing Dynasty which caused massive immigration. This indigenous group was also used as a special force by the CIA in Guerrilla units in the 1960s. What indigenous group is this?

Question Answering System	Guess
https://chat.openai.com/chat	Couldn't guess
https://you.com/	Couldn't guess

Answer: Hmong Indigenous group

Techniques used:
Multihop, oblique paraphrasing, Misleading

Why is this question difficult and interesting?:

The question posed to ChatGPT and You.com presented a challenging task as it required understanding broad facts about a specific indigenous group without any additional context or details. This made the question difficult and interesting, as it required the AI systems to rely solely on their knowledge base and reasoning skills to identify the correct answer.

My question is correct as it asks for a specific indigenous group without providing any additional context or details that could potentially give away the answer. The mention of China and its connection to the Hmong people was included to provide a specific but disconnected fact that could aid in the identification of the group.

Adversarial writing techniques were used in this question by presenting specific but disconnected facts about the indigenous group that could potentially lead the AI systems astray. By avoiding the well-known presence of the Hmong people in Southeast Asia and instead focusing on their population in China, the question becomes more challenging and specific.

The question answering systems were unable to produce an answer due to the lack of context and specific details provided in the question. Without additional information, it becomes difficult for the systems to accurately identify the correct answer.

The question is pyramidal because it starts with a broad topic (an indigenous group) and gradually adds more specific but disconnected details (their population in China) to narrow down the search space and make the question more targeted. However, without additional information, the question remains challenging and difficult to answer.

Question: This city is known as Asia's only Latin city. In 1569, it was chosen as a site for Spanish resettlement and became the official headquarters of the Spanish on approval by King Philip IV of Spain. Currently, this city is known as the 'sardine capital' of the country it is located in. Name the country this city is located in.

Non Pyramidal Question: Name the country where the city, also known as Asia's only Latin city, the site of Spanish resettlement in 1569 and the Sardine capital of its nation is located.

Category: Geography

Source:
https://en.wikipedia.org/wiki/Zamboanga_City

Question Answering System	Guess
ChatGPT	Philippines
You AI	Philippines

Answer: Philippines

Techniques Used: Multi-step, Novel Clues

Why is the Question Interesting and difficult?

This question talks about the only Latin city in Asia. In this question, we attempt to ask a question answering system about the location of this city. The facts stated in this question are all available on the city's Wikipedia page. This question is interesting because it has to do with Latin influence in Asia. Different parts of South-East Asia, were at different times, under the rule of the British, Portuguese and the Dutch and these parts of Asia have imbibed some of the cultural aspects of their invaders. However, no country as such, has a 'British city' or a 'Dutch city'. Zamboanga city is one of a kind.

This question is difficult because we try to trick the computer into thinking that we're going to ask about the name of the city when in fact, we are going to ask about the city's location. In addition to this, the question is pyramidal and doesn't narrow down on what's being asked until the end of the question while progressively giving more information about the city through novel clues.

Both ChatGPT and You AI are able to answer this question presumably because they are able to use the clues to correctly identify the city.

Question: This stateless minority Muslim group has been subjected to several military crackdowns in the past 50 years and is also known as one of the most persecuted minority groups in the world. They cannot obtain citizenship from the country they claim to originate from because that country doesn't recognize their community as one of its national races. What is the old name of the country this community claims to originate from?

Non Pyramidal Question: What is the old name of the country from which the group which has been subjected to several military crackdowns, cannot obtain citizenship of the country in question, and is known as the most persecuted minority in the world, claims to originate from?

Question Answering System	Guess
ChatGPT	Burma
You AI	Burma

Answer: Burma

Why is the Question Interesting and difficult?

This question talks about the Rohingya community which is known to be one of the most persecuted minorities in the world because this community isn't recognized by any sovereign state in the region it occupies.

This question is difficult because we try to trick the computer into thinking that we're going to ask about the name of the community, when in fact, we are going to ask it about the old name of the country that this community claims to originate from.

This question is interesting because it talks about a community that is, according to the UN officials and the Human Rights Watch, a victim of ethnic cleansing. Probes by the UN have found evidence of increasing incitement of hatred and religious intolerance by "ultra-nationalist Buddhists" against Rohingyas while the Myanmar security forces have been conducting "summary executions, enforced disappearances, arbitrary arrests and detentions, torture and ill-treatment, and forced labour" against the community. Many people around the world are unaware of the atrocities that the Rohingyas were victims

History

Question: This country celebrates 'Tet Nguyen Dan' which marks the beginning of the lunar calendar year. Several decades ago, a western military entered this country with the purpose of preventing communal takeover of the region. It was engaged in a battle so horrific that the western soldiers who fought there were said to have been 'ground like meat'. This battle's name has a popular western food item in its name. Name the food item.

Non-Pyramidal question: Name food item present in the name of the battle, between the country that celebrates 'Tet Nguyen Dan' and a western power, that was so horrific that the western soldiers who fought there were said to have been 'ground like meat'.

Category: History

Techniques Used: Multi-step, Novel Clues, Common Link

Source:
https://en.wikipedia.org/wiki/Battle_of_Hamburger_Hill

Question Answering System	Guess
ChatGPT	Hamburger
You AI	Hamburger

Answer: Hamburger

Why is the Question Interesting and difficult?

This question talks about the Battle of Hamburger Hill, between Vietnam and the United States. In this question, we attempt to ask a question answering system about the name of the food item referenced in the name of the battle. The facts stated in this question are all available on the battle's Wikipedia page.

This question is difficult because we try to trick the computer into thinking that we're going to ask about the name of the battle when in fact, we are going to ask about the name of the food item referenced by the battle. In addition to this, the question is pyramidal and doesn't narrow down on what's being asked until the end of the question. It uses the common link between the Battle of Hamburger Hill and the Western love for hamburgers to create a question.

This question is interesting because it has to do with the Vietnam war. Official histories of the engagement refer to it as Hill 937 after the elevation displayed on US Army maps, but the US soldiers who fought there dubbed it "Hamburger Hill," suggesting that those who fought on the hill were "ground up like hamburger meat" in grim reference to the Battle of Pork Chop Hill during the Korean War.

Both ChatGPT and You AI are able to answer this question presumably because they are able to use the clues to correctly identify the name of the food item that is also present in the name of the described battle.

Question: This gemstone is commonly confused with another gemstone that is known to be one of the inspirations for the first full length English detective novel. It was once owned by the Delhi Sultanate and later a part of the Mughal emperor Shah Jahan's 'peacock throne'. What is this gemstone's name?

Non-pyramidal question: Name the gemstone that is confused with the inspiration for the first full length English

detective novel, was owned by the Delhi Sultanate and later a part of Shah Jahan's throne.

Category: History

Source:

https://en.wikipedia.org/wiki/Daria-i-Noor

https://en.wikipedia.org/wiki/Koh-i-Noor

Question Answering System	Guess
ChatGPT	Koh-i-Noor
You AI	Koh-i-Noor

Answer: Daria-i-Noor

Techniques Used: Multi-step, Novel Clues

Why is the Question Interesting and difficult?

The question talks about the gemstone Daria – i – Noor. In this question, we attempt to ask a question answering system the name of this gemstone. The facts stated in this question are all available on the ts Wikipedia page.

This question is difficult because we try to trick the computer into thinking that we're going to ask about the Kohinoor diamond when in fact, we are going to ask about the Daria-i-noor diamond. In addition to this, the question is pyramidal and doesn't narrow **Question:** This group fought a Civil War against the government of an island nation along the coast of India for the creation of a separate state in the north east of the country. In 1991, who was assassinated by a woman belonging to this group?

Non Pyramidal Question: Who was assassinated in 1991 by the group who fought against the government of an island nation along the coast of India in the civil war?

Category: History

down on what's being asked until the end of the question.

This question is interesting because it references the Kohinoor diamond which is known to be the inspiration for the first full length English detective novel. Its ownership has been disputed by many countries and is a source of great intrigue. The Daria-i-noor diamond is said to have been found in a mine in Andhra Pradesh, India. It was once owned by the Delhi Sultanate and later the Mughal Emperor Shah Jahan. It is commonly confused with the Kohinoor due its structural and cultural similarities.

Just as the facts in the question stated, ChatGPT and You AI confused Daria-i-Noor for Koh-i-Noor. This could be because the history of India is seen as so closely connected with the Koh-i-Noor that the answering systems may not be exposed to much information about the less discussed Daria-i-Noor. Therefore, when the answering systems see 'gemstone' and India related history, it defaults to answering Koh-i-Noor.

Source:

https://en.wikipedia.org/wiki/Sri_Lankan_Civil_War

https://en.wikipedia.org/wiki/Liberation_Tigers_of_Tamil_Eelam

Question Answering System	Guess
ChatGPT	Rajiv Gandhi

Bard	Rajiv Gandhi

Answer: Rajiv Gandhi

Techniques Used: Oblique Paraphrasing, Multihop

Why is the Question Interesting and difficult?

This question talks about Sri Lanka and the civil war that was fought in the country for the major part of its history since its Independence. During the civil war, they had also caused unrest in the neighboring country, India's state of Tamil Nadu by assassinating the then PM of India, Rajiv Gandhi. This question is interesting because not many people know about the country other than the fact that it is an island in the Indian Ocean.

The question is difficult to answer because we use oblique paraphrasing in mentioning the Sri Lankan civil war vaguely by talking about a civil war in an island nation and then hoping to ask about who was assassinated as a consequence of this war. This question also is pyramidal because we ask about the concrete answer in the end.

The question is also diverse because it talks about a country with a small population but rich history, often not talked about in mainstream media. Both ChatGPT and Bard are able to get the question right because there is no other island nation along the Indian coast that fought a civil war apart from Sri Lanka and only one person was assassinated by this group in 1991.

Question:
This person has been the prime minister of a country and rishi sunak is currently the prime minister of this same country. There were two major parties in the United Kingdom in the 1900s and he was not part of the conservative party. He was the last prime minister of the other party. This party was founded by John Russell. Who is this person?

Non-Pyramidal Question: Name the prime minister of united Kingdom from 1916 to 1922

Category: History

Source:
https://en.wikipedia.org/wiki/David_Lloyd_George

Question Answering System:	
You	Clement Attlee
ChatGPT	Clement Attlee

Answer: David Lloyd George

Techniques Used: Common link, current events and oblique paraphrasing

Why is the question interesting and difficult:

This question is difficult because it uses a common link to state that the person asked in the question and rishi sunak were the prime ministers of the same country but during different time frames, hence the QA system has to understand which country they were a prime minister of, to narrow down the possible prime ministers that have been prime minister previously since rishi sunak is the current prime minister. Moreover since, these QA systems are still not updated that rishi sunak is the current prime minister, it makes it more difficult for the system to identify the country in order to find the other prime minister. This question also uses oblique paraphrasing by implicitly saying that this person was the prime

minister for the other party which was founded by John Russell. This confused the QA system because it thought that the question was referring to the labor party instead of the liberal party. However the QA system misinterpreted it as the Labour party and this party was not founded by John Russell hence the QA understood and predicted the party wrongly even when the necessary information was given.

Additionally since the question uses multihop to incorporate different details about this person, it is more difficult for the QA system to understand all the details and match it to the identity of a person.

This question is correct because it has the accurate factual information quoted from reliable sources. Additionally, this question is pyramidal because it provides a series of hints starting with very broad information such as providing details about his career and gradually narrowing down to specific details by stating the party that he was a prime minister of and specifying the name of the person who founded the party.

Question: This pagoda has been rebuilt countlessly due to it being destroyed and burnt in wars by opponents. The story behind the pagoda is that one day an emperor lingered for a son and dreamt that the Goddess of Mercy was sitting on a lotus flower offering him a son. To pay his respect to the goddess, he ordered the construction of a small pagoda in the form of a lotus, overlooking a pond. Name the country this structure is located in.

Non Pyramidal Question: Name the country that is home to the pagoda was continuously rebuilt and destroyed due to wars and was built by an emperor who longed for a son.

Category: History

Source:

https://en.wikipedia.org/wiki/Vietnamese_art

Question Answering System	Guess
ChatGPT	Vietnam
You AI	China

Answer: Vietnam

Techniques Used: Multi-step, Novel Clues

Why is the Question Interesting and difficult?

This question talks about the One Pillar Pagoda in Hanoi, Vietnam. All information can be found on its Wikipedia page.

This question is difficult because we try to trick the computer into thinking that we're going to ask about the name of the Pagoda, when instead we ask it about the country it is located in. This question utilizes the multi-hop technique to confuse the computer. The question is pyramidal because it provides a story as a clue.

This question is interesting because it has to do with Vietnamese art and culture.

Furthermore, this structure is one of the most ancient structures in Hanoi and has an interesting story behind its construction too.

Only ChatGPt was able to answer this question correctly, whereas You AI says that the clues that the story provides match the Nine-Story Pagoda, which is located in China. Like this One Pillar Pagoda, the Nine-Story Pagoda has a long history of being destroyed and reconstructed due to warfare and other disasters. The structure is known for its unique, lotus flower-inspired design, which includes nine tiers or stories that are decorated with intricate carvings and other artwork. This means that it couldn't correctly match the story to the monument.

Science

Question: What was the space mission successfully completed by a national space organisation reaching the orbit of a nearby planet at a cost less than the budget of the space science fiction thriller, Gravity?

Non Pyramidal Question: What was the mission completed by ISRO at a cost less than the budget of the film gravity?

Category: Science

Source:
https://en.wikipedia.org/wiki/Mars_Orbiter_Mission

https://www.nbcnews.com/science/space/why-indias-mars-orbiter-mission-cost-less-gravity-movie-n210681
https://en.wikipedia.org/wiki/Gravity_(2013_film)

Question Answering System	Guess
ChatGPT	Mars Orbiter Mission

Bard	Mars Orbiter Mission

Answer: Mars Orbiter Mission/Mangalyaan

Techniques Used: Common Link, Logical Reasoning

Why is this question interesting and difficult?

This question is interesting because it asks about the Indian Space Organization's Mission that was completed in less than the movie Gravity's budget. This accolade went to the Mars Orbiter Mission for having

completed in 74mn USD while Gravity spent around 100mn USD on its movie.

The question employs logical reasoning and common link to compare the budgets of a movie and the space mission by ISRO in a pyramidal question as there have been multiple successful space missions by national space organizations to a nearby planet but only one that cost less than the movie Gravity.

ChatGPT and Bard have gotten this right because there's an article which both might have learned from and seem to have an n gram match that helps them get the question right despite all the techniques employed.

Question: Imagine you lived in UK in the 1970s, if you noticed a loud explosion like sound everytime a flight passed by, what kind of flight is passing by?

Non-Pyramidal Question: What flight would create a loud explosion like noise when it flies?

Category: Science

Source:
https://en.wikipedia.org/wiki/Sonic_boom

https://en.wikipedia.org/wiki/Concorde

Question Answering System	Guess
ChatGPT	Concorde
Bard	Concorde

Answer: Concorde

Techniques Used: Common Sense, Multihop

Why is the question interesting and difficult?

This question is interesting because it talks about commercial supersonic flights which were around in the late 70's till early 2000's.

Concorde flights enabled commercial supersonic travel and were quite fast compared to todays flights. However, after an accident in early 2000's it lost its popularity to more efficient and larger flights made by boeing and airbus and couldn't justify the cost amidst the question of safety from the accidents.

The question is difficult because it talks about a scenario where you hear a loud bang when a flight passes by. This is a question of common sense and science that when a loud thud or bang is heard it is a supersonic travel and the sonic boom associated with it. This was achieved by the concorde commercially in late 70's. It also employs multihop by giving a hypothetical scenario and then moving on to asking the real question.

However, the difficulty of the question doesn't defeat GPT and Bard, as they're able to identify the phenomenon as sonic boom and identify that concorde was a commercial supersonic aircraft linking them both.

Question:

There existed a diplomat from Vietnam. Although he was lauded for his scientific achievements, he was a strong believer in eastern philosophy - he was a pioneer in chemical carcinogenesis where it involved the fundamental ring systems and derivatives of 1,2-benzanthracene, the dibenzopyrenes, steranthrenes, anthanthrene, 1,2,3,4-dibenzanthracene, a variety of benzo and dibenzofluoranthenes. He was also a secular buddhist who became an elected president with thousands of workers stranded in France in the 1950s. He also travelled to the United States and also to Cambodia, to meet the pope of Cao Đài and his entourage in Cambodia. Who is this man?

Non-Pyramidal Question:
There existed a diplomat from Vietnam. Although he was lauded for his scientific achievements, he was a strong believer in eastern philosophy - he was a pioneer in chemical carcinogenesis, and he was also a secular buddhist. Who is this man?

Category: Science

Source:
https://en.wikipedia.org/wiki/B%E1%BB%ADu_H%E1%BB%99i

Question Answering System:	

You	**Nguyễn Ký**
ChatGPT	**Nguyễn Sinh Cung**

Answer: Bửu Hội

Why is the question interesting and difficult:

The question about the influential people involved in Vietnam is interesting and difficult because it incorporates various traits that are familiar to them, such as their travels to France and the US and their meeting with the Pope of Cao Đài. The question is crafted to specifically highlight the chemicals associated with these people to give a clear picture of who is being referred to. The adversarial writing technique is used to intentionally confuse the AI systems by providing vague details and multiple possible answers. The question answering system may have struggled to perform because it lacked the nuanced understanding and context necessary to accurately identify the person in question. The question is pyramidal in nature because it starts with general information and gradually becomes more specific in order to provide a more complex answer.

Techniques used:
Misleading, oblique phrasing

Question:

```
Who was a Chinese
mathematician who was
posthumously award the State
Natural Science Award in the
1980s, as well as having his
theorems completed by Luc
Teirlinck, as well as solving a
generalised      Kirkman's
schoolgirl   problem   and
submitting it to 数学通报 (Acta
Mathematica   Sinica,   not
included in the question) and
privately engaged in research?
```

Category: Science

Source:
https://en.wikipedia.org/wiki/Lu_Jiaxi_(mathematician)

Question Answering System	Guess
https://chat.openai.com/chat	Sun Jiaguang
https://you.com/	Chen Jingrun

Answer: Lu Jia Xi

Techniques used:
Oblique paraphrasing, Misleading

Why is this question difficult and interesting?:

The question posed to the AI systems, which involved using the Chinese characters for Acta Mathematica Sinica, was challenging due to its specificity and potential for confusion. This made the question interesting, as it required the systems to understand and differentiate between Chinese mathematicians who submitted their papers to this particular journal.

My question was correct because it specifically asked for a living mathematician who had published in the journal, rather than a deceased mathematician who had been awarded the State Natural Science Award.

Adversarial writing techniques were not explicitly used in this question, although the specific details and context of the question could potentially have led the AI systems to focus on certain keywords or misinterpret the intended meaning.

Question: This is an organic hydrocarbon is a byproduct of incomplete combustion of many materials. This is colorless and it is in the same state of water at room temperature. It is an aromatic hydrocarbon because of its cyclic continuous pi bonds between carbon atoms.The bonds between carbons are stronger than a single bond but weaker than a double bond due to electron delocalisation. What chemical compound is this?

Non-Pyramidal Question: Name the compound which has the molecular formula C_6H_6.

Category: Science

The question answering systems performed the way they did because they relied heavily on the keywords in the question and did not take into account important contextual information such as whether the mathematician was alive or deceased. You.com returned an answer related to Chen Jingrun, a deceased mathematician who had been awarded the State Natural Science Award, while ChatGPT returned an answer related to Sun Jiaguang, a living mathematician who had published in the journal but had not been awarded any posthumous awards.

The question is pyramidal because it starts with a specific detail (using the Chinese characters for Acta Mathematica Sinica) and gradually adds more context and information (looking for a living mathematician who published in the journal) to narrow down the search space and make the question more targeted. This allows for a more precise and accurate answer. More information can be found here: https://www.thss.tsinghua.edu.cn/en/faculty/jiaguangsun.htm

Source:
https://pubchem.ncbi.nlm.nih.gov/compound/Benzene

Question Answering System:	
You	Benzene
ChatGPT	Benzene

Answer: Benzene

Techniques used:
Multihop, oblique paraphrasing, Misleading

Why is the question interesting and difficult:

This question is difficult because multihop is used and the question gives information about different details, for instance it gives details about the structure, and uses oblique paraphrasing to indirectly say what is its state at room temperature, and talks about which reaction produces this compound.

The part of the question that states that this compound is a byproduct is a bit misleading because there might be numerous byproducts formed during the incomplete combustion. Hence leading the GPT to focus on other products that are organic hydrocarbons and could fit the description, resulting in a wrong guess. However, the final hint that states about the bond strength between the carbon atoms would definitely allow humans to guess that an organic hydrocarbon that is aromatic and is a liquid at room temperature with electrons delocalised into the carbon ring is benzene. However for a GPT, putting together all this information and understanding that the electrons are delocalized into the benzene ring leading to a carbon bond weaker than double bond but stronger than single bond would be more tough.

This question is interesting because benzene is not the first compound that comes to mind usually when we think of byproducts of incomplete combustion, carbon monoxide is more common, hence incorporating that information to this question along with the structure of benzene makes this question unique.

This question is correct because it has the accurate factual information quoted from reliable sources. Additionally, this question is pyramidal because it provides a series of hints starting with very broad information such as stating that this product is an organic hydrocarbon that is formed as a byproduct during incomplete combustion. There are numerous compounds that satisfy this requirement hence it does not allow the QA system to narrow down to a smaller set of possible answers. Thereafter it gradually narrows down to specific details by stating the hydrocarbon's structure.

TV and film

Question:

This film has been entirely composed, directed and produced by a single person, for the most part. In this film, it stars a Filipino driver who is fascinated by space travel. Using elements of repetition and flashbacks, as well as covering controversial topics such as circumcision, this Filipino man finds solace in thinking that success is achievable in the United States. Upon reaching here, he realises that success is found "within". What film is this?

Non-Pyramidal Question: This movie stars a Filipino jeepney driver who is fascinated by space travel, where he dreams of moving out of his village to settle in the United States to make his dream of being an astronaut come true. What film is this?

Category: TV and Film

Source:
https://en.wikipedia.org/wiki/Perfumed_Nightmare
https://mubi.com/films/the-perfumed-nightmare

Question	

Answering System:	
You	**Couldn't guess**
ChatGPT	**Anak ng Tinapa**

Answer: Mababangong Bangungot, or "perfumed nightmare"

Techniques Used:
Multihop, Oblique Phrasing

Why is the question interesting and difficult:

This question presents an interesting challenge as it involves a relatively obscure film that is not well-known in the Asian film industry. Additionally, the question includes controversial elements and mentions multiple countries, which could potentially confuse the AI systems. Despite these challenges, the question is correct and accurately describes the film in question. The adversarial writing techniques used in the question include vagueness and obfuscation, which make it difficult for the AI systems to accurately identify the film. The question answering system may have struggled to perform because the film in question is not well-known and therefore may not be indexed in the databases that the AI systems use. Finally, the question is pyramidal in nature, as it presents multiple layers of complexity and difficulty that must be overcome in order to arrive at the correct answer.

Question: For more than a decade, this movie was never officially released in the US or Canada except for the purpose of film festivals due to its 'crude' nature. In this movie, the government, in order to control the unemployed youth, votes on a law, which states that every year, students from a randomly chosen class will be transferred to a secluded island where they will have to fight to the last person standing. Name the novel this movie was based on.

Non Pyramidal Question: Name the novel which was the source of inspiration for the movie that was never officially in the US or Canada except for the purpose of film festivals due to its depiction of a story where , the government, in order to control the unemployed youth, states that every year, students from a randomly chosen class will be transferred to a secluded island where they will have to fight to the last person standing.

Category: TV and Film

Techniques Used: Multi-step, Novel Clues

Source:
https://en.wikipedia.org/wiki/Battle_Royale_(film)

Question Answering System	Guess
ChatGPT	Battle Royale
You AI	Battle Royale

Answer: Battle Royale

Why is the Question Interesting and difficult?

This question talks about the 2000 action thriller 'Battle Royale', in which the government tries to take a 'survival of the fittest' to deal with the after effects of raging unemployment. The facts related to this question are available on its Wikipedia page.

This question is difficult because we try to trick the computer into thinking that we're going to ask about the name of the movie, when in fact, we are going to ask it about the novel that was the source of inspiration for this movie.

This question is interesting because it talks about a movie that wasn't well received in the

US due to its apparent 'crude' and 'graphic' storyline. This film was screened to a test audience in the US during the early 2000s, soon after the Columbine High School massacre, which resulted in a negative reaction to the film's content. This movie is also said to have influenced the reception of The Matrix (1999) because much of the test audience for Battle Royale condemned the film for its 'mindless' and gratuitous violence. In addition to this, the novel on which this movie was based shares its name with the movie.

Both ChatGPT and You AI are able to answer this question presumably because they are able to use the clues to correctly identify the name of the movie.

Question: Who was born in Shanghai in the 1930s, who subsequently moved to Macau and then to Hong Kong - where the person then joined the Shaw Brothers in the 1960s to develop multiple films under his belt with the funding of the Shaw brothers, as well as being a honorary member of Photography Societies in Singapore, Argentina and Brazil due to his heavy influence in film and photography?

Category: TV and Film

Source:
https://en.wikipedia.org/wiki/Fan_Ho

Answer: Fan Ho

Question Answering System	Guess

https://chat.openai.com/chat	Raymond Chow
https://you.com/	Sammo Hung Kam-bo

Techniques used:
Multihop, oblique paraphrasing, Misleading

Why is this question difficult and interesting?:
The question of identifying a specific person who joined Shaw Brothers, a prominent Hong Kong film production company, presented a challenge for the question answering systems. The task required understanding the nuances of the question, as the person was not one of the Shaw brothers themselves but had an alliance with the company. This level of complexity makes the question interesting and difficult, as it requires understanding contextual information beyond just the keywords.

My question is correct because it specifically asks for a person who joined the Shaw Brothers, rather than

information about the co-founders or anyone else associated with the company. It also provides additional details such as the individual's birthplace and migration, which are important for identifying the correct answer.

In the question, adversarial writing techniques were used by including additional details that could potentially mislead the question answering system into providing an incorrect answer. For example, mentioning that the person was not originally part of the Shaw family could potentially lead the system to exclude the co-founders from consideration, despite their close association with the company.

The question answering system likely struggled to understand the nuances of the question and instead focused solely on the keyword "Shaw Brothers," leading it to provide an answer related to the co-founders rather than the specific person in question. The answer provided by You.com was similarly

Question: This film originates from the country known to be the home of the world's largest fish market. It follows a theater director, who directs a multilingual production of a play written by Anton Chekhov, while dealing with the death of his wife. This film became the first film in 13 years to win an Oscar from its country. Name the film.

Non-pyramidal question: Name the film, which is about a theater director, who directs a multilingual production of a play written by Anton Chekhov, while dealing with the death of his wife and originates from the country home to the world's largest fish market, and became the first film in 13 years to win an Oscar from its country.

Category: TV and film

Source:

inadequate, as it ignored important details about Fan Ho's alliance with the company. This highlights the difficulty of developing robust question answering systems that can accurately understand and respond to complex questions.

The question is pyramidal because it includes additional details and context that narrow down the search space and make it more specific. Starting with a broad keyword like "Shaw Brothers" and gradually adding more information about the specific person and their background allows for a more targeted and precise question.

This question is particularly confusing as the 1900s were considered the Golden Era in East-Asian filmmaking - leading to the rise of many well-established actors and filmmakers within this era. It makes it particularly difficult for these AI systems to pinpoint the correct person in question.

https://en.wikipedia.org/wiki/List_of_Japanese_Academy_Award_winners_and_nominees

https://en.wikipedia.org/wiki/Drive_My_Car_(film)

Question Answering System	Guess
ChatGPT	Departures
You AI	Departures

Answer: Drive My Car

Techniques Used: Multi-step, Novel Clues

Why is the Question Interesting and difficult?

This question talks about the film 'Drive My Car'. In this question, we attempt to ask a question answering system about the name of the film by tricking it into thinking that we're asking about the country of origin of the film. The facts stated in this question are all available on the film's Wikipedia page.

This question is difficult because it hides the critical clues using novel phrasing. For example, the play written by Anton Chekhov is 'Uncle Vanya' and the home to the largest fish market is Japan.

This question is interesting because it has to do with the first film from Japan. to win an Oscar after thirteen years.

Both ChatGPT and You AI completely get the reasoning wrong and hence they get the answer wrong. This is probably because of the way the clues are ordered. Both systems are not able to identify the movie from the plot and subsequently cannot identify the first film in 13 years, from Japan, to win an Oscar.

Question:

The actors in this movie and this movie received multiple oscar awards this year. They received awards for best picture, best director, best original score, best actress, best original screenplay and more. The lead female actor is originally from Malaysia. She also acted in another movie where she played the role of the mother for the lead male actor and acted together with Constance Wu and awkwafina. Who is this female actor?

Non-Pyramidal Question: Name the lead female actor from Everything Everywhere all at once

Category: TV and film

Source:
https://en.wikipedia.org/wiki/Academy_Awards

https://www.imdb.com/name/nm0000706/
https://www.imdb.com/title/tt3104988/

Question Answering System:	Guess
You	Yuh-Jung Youn
ChatGPT	Michelle Yeoh

Answer: Michelle Yeoh

Techniques Used: Misdirection, Novel Clues

Why is the question interesting and difficult:
This question is difficult because the question misdirects the QA system into thinking that the question is asking about the movie that won numerous oscar awards, instead of the lead female actor of the movie. Additionally, since the some of the QA systems are not up

to date and they only have learnt from data till 2021, they make an incorrect guess of what movie is being referred to because it guesses a movie that received many oscars in the year 2021 instead of 2023, although the question specifies that the movie received the awards this year.

Literature

Question: Known for having ruled his kingdom during the golden age of telugu literature, what is the famous line that the emperor wrote to praise the language that his work amukthamalayada was written in?

Non Pyramidal Question: What was the famous line of praise about telugu in the work amukthamalayada?

Category: Literature

Source:
https://en.wikipedia.org/wiki/Krishnadevaraya#Telugu_literature

https://en.wikipedia.org/wiki/Amuktamalyada

Question Answering System	Guess

ChatGPT	Desha bhashalandu Telugu Lessa
Bard	Telugu, the language of the gods, is sweet to the ear and pleasing to the mind

Answer: Desha bhashalandhu Telugu Lessa

Techniques Used: Multilingual, Common Link, Multihop

Why is this question interesting and difficult?

This question is interesting because it talks about a language and its golden era of its literature. Telugu is a dravidian language spoken mainly in South India by around a 100 million people in the telugu states of Andhra Pradesh and Telangana. Recently, in a survey

conducted in the United States it was found that this language was the fastest growing language in terms of speakers in the United States from 2010-2017.

The question asks about a sentence that was written by King Krishna Deva Raya of the Vijayanagara Empire praising the language. The question employs a multilingual technique along with multihop and common

link to piece together information about the king who ruled during the golden age of telugu literature and the work he wrote praising the language.

ChatGPT gets the answer right however, bard comes close by answering the question in a translation into english. Therefore, it seems that multilingual questions are harder for bard to answer.

Question:
There is a script from an ancient classical language that was spread in the 12th century to Southeast Asian countries. Upon modifying this script with tonal markers, it gave birth to languages like Thai, Javanese and Balinese. This script is also tied to dynasties in South Asia during the same period. What is the ancient script in question?

Non-Pyramidal Question:
There is a script from an ancient classical language that was spread in the 12th century to Southeast Asian countries. Upon modifying this script with tonal markers, it gave birth to languages like Thai, Javanese and Balinese. What is the ancient script in question?

Category: Literature

Source:
https://en.wikipedia.org/wiki/Thai_script
https://scroll.in/magazine/1007954/the-journey-of-pallava-script-from-tamil-nadu-to-south-east-asia

https://en.wikipedia.org/wiki/History_of_Indian_influence_on_Southeast_Asia

Question Answering System:	
You	**Brahmi Script**
ChatGPT	**Brahmi Script**

Answer: Pallava Script

Techniques Used:
Oblique Phrasing, Misleading

Why is the question interesting and difficult:

The question of identifying the script used to write ancient classical languages in South Asia is an interesting and difficult one. The ancient South Asian scripts have undergone many mutations over the centuries, making it challenging for the AI systems to pinpoint the specific script. The question further adds complexity by involving the Brahmi script, which gave birth to the Pallava script to write ancient classical languages like Tamil. While

the AI systems partially answered the question correctly by suggesting the Brahmi script, a more nuanced understanding is needed to answer the question fully. The Pallava script would be a more distinct and correct answer due to its development in the Pallava Language and its connection to maritime trade, which led to the Indianisation of Southeast Asia, primarily by the South Indian Kingdoms at that time. Thus,

Question:
There is a writing system developed in this country in question, where it was created in opposition to the writing system used by the educated class. It is written in morpho-syllabic blocks, created to combat the complicated Chinese characters and writing system. It is also agglutinative and it was made official after independence from the Japanese. It also follows a writing style similar to the Indic tradition. What language is this?

Non-Pyramidal Question:
There is a writing system developed in this country in question, where it was created in opposition to the writing system used by the educated class. It is written in morpho-syllabic blocks, created to combat the complicated Chinese characters and writing system. It also follows a writing style similar to the Indic tradition. What language is this?

Category: Literature

Source:
https://en.wikipedia.org/wiki/Hangul

answering this question requires a deep understanding of the intricate relationships and histories of South and Southeast Asian cultures and languages.

Question Answering System:	
You	**Korean**
ChatGPT	**Korean**

Answer: Korean

Techniques Used:
Multihop

Why is the question interesting and difficult:

This question is interesting and difficult because it employs adversarial writing techniques to test the ability of AI systems to accurately answer vague questions that span multiple countries and time periods. The question provides few details about the language in question, further complicating the task for AI systems.

Despite these difficulties, the AI systems were able to accurately identify the language in question. The pyramidal structure of the question, with its intentional vagueness and use of multiple countries, tested the robustness of the AI systems and their ability

to identify key pieces of information to arrive at the correct answer. The success of the AI systems in this task demonstrates the power of modern language models and their ability to handle complex and challenging natural language tasks.

Question:
This person wrote the play for one of the greatest love stories of all time. He was born in England and was known as that country's poet. He wrote another play which is famous for one of the lines used in the play, "pound of flesh". Name this other play.

Non-Pyramidal Question: Name the play that is known for this famous line, "Hath not a Jew eyes?".

Category: Literature

Source:

Question Answering System:	
You	Merchant of Venice
ChatGPT	Merchant of Venice

Answer: Merchant of Venice

Techniques Used: Novel Clues, Misdirection, Multi- Step

Why is the question interesting and difficult:

This question is difficult because it uses misdirection to confuse the QA system into thinking that the question is going to ask the identity of the person who wrote the play which is the one of the greatest of all time love stories, since the question starts off by describing the person.

The question also uses oblique paraphrasing by using the phrase greatest loves stories of all time instead of specifying that the play was Romeo and Juliet, so as to confuse the QA system into thinking about other famous love stories, for instance, 'Waiting For Godot' which was written by Samuel Beckett. The question also references a specific line from the play, which might be difficult for the system to retrieve, compare and choose the correct play that uses this particular line.

This question is interesting because it also uses multihop to connect different types of information that provide details about the play. For example, it introduces the name of the person who wrote the play, by stating

where he was born and describing details about another famous play.

Question: The president of one of the Asian countries is known to be an avid follower of this philosopher. This philosopher believes that the welfare of a country depended on the moral cultivation of its people, beginning from the nation's leadership. This philosopher is?.

Non Pyramidal Question: Name the philosopher who is a source of inspiration to a President of an Asian country and believes that the welfare of a country depended on the moral cultivation of its people, beginning from the nation's leadership.

Category: Literature

Techniques Used: Multi-step, Novel Clues

Source: https://www.economist.com/the-economist-explains/2021/06/23/how-did-confucianism-win-back-the-chinese-communist-party

Question Answering System	Guess
ChatGPT	Confucius
You AI	Confucius

Answer: Confucius

Why is the Question Interesting and difficult?

This question talks about Confucius, the Chinese philosopher and scholar. The facts stated in this question are all available on the philosopher's Wikipedia page

This question is difficult because it is not apparent which Asian country is being talked about. In addition to this, the question gives additional information about the subject's beliefs, but it is still hard to determine which philosopher it is.Confucius is a highly respected philosopher across South-east Asia. Several Asian governments are known to use 'filial piety' and other such Confucian values to garner support for themselves.

This question is interesting because the President of China, Xi Jinping is known to lean on Confucianism to gain support for himself and the Chinese Communist Party. In addition to this, Confucius is known to be the inspiration for the 'Analects' which has been one of the most widely-read and studied books in China for the last 2,000 years, and continues to have a substantial influence on Chinese and East Asian thought and values today.

Both ChatGPT and You AI are able to answer this question presumably because they are able to use the clues to correctly identify the philosopher.

Others

Question: This person is the current prime minister of the country that consists of multiple countries, where one of them is the birthplace of Shakespeare. He was born in a port city that is located in Hampshire, England. Who is this person?

Non-Pyramidal Question:
Name the person who is the current prime minister of the United Kingdom.

Category: Diversity

Source:
https://www.britannica.com/biography/Rishi -Sunak

Answer: Rishi Sunak

Question Answering System:	
You	Boris Johnson
ChatGPT	Boris Johnson

Techniques: Current Events, Multistep, Oblique Paraphrasing

Why this question is difficult:

This question is difficult because some of the GPT models are not being up to date, hence they are not aware of the current events. For instance in this question, we asked about the current prime minister of United kingdom, however the data that it has last stored in the GPT system is the previous prime minister which is Boris Johnson instead of the current prime minister Rishi Sunak, hence it returns the wrong answer, because the GPT still thinks that boris Johnson is still the prime minister which is not the case. This question is correct because it gives accurate details about the prime minister as quoted from reliable sources.

The multihop technique and oblique paraphrasing is used in this question by stating that the country that he is a prime minister of, consists of multiple countries of which one of them is the birthplace of Shakespeare. Hence the question does not explicitly state which country but gives some hints about the country. With the hints given which is the birthplace of Shakespeare, the GPT has to derive that the country is England and the question is asking about the prime minister of the United kingdom. Therefore after making these derivations, the GPT can use the additional information that the prime minister was born in

Hampshire England to make a guess on who the prime minister is. The GPT gave Boris Johnson as the answer, which is wrong. He was the previous prime minister and he was not born in England, he was born in New York instead.

Additionally, this question is pyramidal because it provides a series of hints starting with very broad information such as providing details about the country that the person is a prime minister of and gradually narrowing down to specific details by stating the where the prime minister was born.

Question: This famous monument can be seen in the popular video game Animal Crossing for the Nintendo GameCube and is located in the country that was home to the world's tallest Ferris wheel, until the High Roller was built. Name this country.

Non-Pyramidal Question: Name the country home to the second largest Ferris wheel in the world and also a monument that can be seen in the popular video game Animal Crossing for the Nintendo GameCube.

Category: Diversity

Source:

https://en.wikipedia.org/wiki/Merlion

https://en.wikipedia.org/wiki/Singapore_Flyer

Question Answering System	Guess
ChatGPT	Japan
You AI	Japan

Answer: Singapore

Techniques Used: Multi-step, Novel Clues, Common Link

Why is the Question Interesting and difficult?

This question talks about Singapore which is home to the Merlion and the Singapore Flyer. In this question, we attempt to ask a question answering system about the location of these monuments. The facts stated in this question are all available on the monuments' Wikipedia pages.

This question is difficult because knowing about what appears in a video game on a medium-popularity platform is a very niche thing. Secondly, the world's tallest Ferris wheel is not the High Roller but the Ain Dubai. Both question answering systems got this question wrong presumably because they do not know the history of the highest Ferris wheels.

This question is interesting because it had to do with the location of the Merlion and the Singapore Flyer. Both these monuments, especially the Merlion, have great historical significance for Singaporeans. Second, the

first four main series games of Animal Crossing are among the best-selling video games for their respective consoles.

Both ChatGPT and You AI get this question wrong because they are not exposed to information about the history of Ferris Wheels. According to them, Japan is the answer because it states that the Daikanransha is the Ferris wheel in question.

Question:

In World War II, a mass killing had taken place in the Southeast Asian peninsula. In this mass killing, many Burmese, Thai, Tamil Indians had been sentenced to the "Death Railway" located between modern day Thailand and Myanmar, to work to their deaths. On top of this, the Japanese Kempeitai had been targeting the Chinese living in the country in question. This country has been Britain's strategic asset and its loss of this country remains the largest surrender to the Japanese. What country is this?

Non-Pyramidal Question: In World War II, the Japanese Kempeitai had been targeting the Chinese living in the country in question. This country has been Britain's strategic asset and its loss of this country remains the largest surrender to the Japanese. What country is this?

Category: History

Source:
https://en.wikipedia.org/wiki/Sook_Ching

Question Answering System:	
You	Thailand
ChatGPT	Thailand

Answer: Singapore

Techniques Used: Multihop, misleading

Why is the question interesting and difficult:

This question is designed to confuse the AI systems by strategically using Thailand. The question includes multiple countries and imperial powers to obscure the correct answer, and the AI systems were unable to identify the specific country in question. Interestingly, Thailand was given as a possible answer despite never having been colonized, highlighting the challenge of using broad geographic references.

Question: What conference did the US President Joe Biden, UK Prime Minister Rishi Sunak and Indian Prime Minister Narenda Modi and other leaders attend in the hindu majority island of the largest muslim populous country in November 2022?

Non Pyramidal Question: What was the conference attended by Joe Biden, and other leaders in Bali in November 2022?

Category: Politics

Source:
https://en.wikipedia.org/wiki/2022_G20_Bali_summit
https://www.g20.org/en/about-g20/#previous-summits

https://en.wikipedia.org/wiki/Islam_in_Indonesia

Question Answering System	Guess
ChatGPT	No answer
Bard	Bali, Indonesia

Answer: G20 Bali Summit

Techniques Used: Novel Clues, Common Link, Oblique Paraphrasing

Why is the Question Interesting and difficult?

This question asks about the G20 Bali Summit that took place in November 2022 which was attended by the leaders of G20 Nations. The United States, UK and India are all part of the G20 group and their leaders had attended the summit that took place in Bali.

This questions uses novel clues because the summit took place in 2022, a common link between the member nations mentioned in the question and oblique paraphrasing to talk about Bali being a hindu majority island in Indonesia, the largest muslim populated country. All these techniques contribute in making it a hard question.

The question is also pyramidal because the UK and the USA are part of the G7 and could've attended a conference as part of the G7 but the mention of India and November 2022 narrows down the group to G20 with a summit that was held in Bali.

ChatGPT and Bard are unable to answer the question correctly because GPT does bad with novel events since its database is limited to information from 2021 and Bard is unable to get it right due to the common link,

however it comes close to guessing the place despite the oblique paraphrase.

Question: After having ruled post Independent India for about 55 years in total under a number of Prime Ministers from its party, who was the first Prime Minister of the country who wasn't from this Party?

Non-Pyramidal Question: Who was the first prime minister of India that didn't belong to the Indian National Congress?

Category: Politics

Source:
https://en.wikipedia.org/wiki/List_of_prime
_ministers_of_India

https://en.wikipedia.org/wiki/Indian_Nation
al_Congress

Question Answering System	Guess
ChatGPT	Morarji Desai
Bard	V.P Singh

Answer: Morarji Desai

Techniques Used: Multihop, Misdirection

Why is this question interesting and difficult?

This question is interesting because it probes about the history of Indian Politics. This is interesting because for the majority of Independent India's history, Indian National Congress reigned the country through its numerous Prime Ministers and repeatedly won consecutive elections. However, the first time a non Indian National Congress Prime Minister came to power was due to Indira Gandhi, the prime minister from INC who imposed a national emergency which garnered opposition from a large part of the population, prompting the people to vote for a party that wasn't Indian National Congress. The first non-INC prime minister was Morarji Desai, from the Janata Party.

This question is difficult to answer for a computer because we employ multihop, common link and misdirection to phrase the question. There were multiple prime ministers in the country from the INC making the base a broad one before we focus on the first non INC prime minister, making the question pyramidal.

ChatGPT gets the question right however, Bard doesn't. This might be due to an n-gram match with Morarji Desai as the first non INC prime minister, and bard comes close with V.P Singh as he was the second non INC prime minister. The misdirection works with bard

Sports

Question: This cricket umpire is known as 'the man with the crooked finger'. He is also famous for his funny way of signaling fours and sixes. Name the umpire.

Non Pyramidal Question: Name the umpire, who due to his funny way of signaling events in cricket, is nicknamed 'the man with the crooked finger'.

Category: Sports

Source:
https://en.wikipedia.org/wiki/Billy_Bowden

Question Answering System	Guess
ChatGPT	David Shepherd
You AI	Billy Bowden

Answer: Billy Bowden

Techniques Used: Multi-step, Novel Clues

Why is the Question Interesting and difficult?

This question talks about the cricket umpire, Billy Bowden who is known for his funny way of signaling fours and sizes. He also holds up a crooked finger to signal 'out', due to which several cricket fans gave him the name 'the man with the crooked finger'.

This question is difficult because the nickname 'the man with the crooked finger', isn't this umpire's first and most well-known nickname. He is more commonly known as the 'dancing umpire'.

A person who watches cricket might find this question interesting because umpiring in cricket requires very standard signals for fours, sixes and wickets. Billy Bowden was an umpire who would signal 'out' with a crooked finger instead of a straight, pointed finger because he has rheumatoid arthritis. He also had a funny, dance-like way of signaling fours and sizes and this also brought him the name 'dancing umpire'. Billy Bowden was known for entertaining the crowd with his signals.

ChatGPT might have answered David Shepherd because he was also famous for his tendency to shake his hand while signaling fours. However, it did not make use of the fact that this umpire is known as 'man with the crooked finger'. In addition, I also found several articles on the internet about a man called David Shepherd who is a member of the Crooked Stick Golf Club. This may or may not have impacted ChatGPT's guess.

Question: These two cricketers share a nickname. One cricketer is the mentor and former player of the team that has won the most Indian Premier League titles. His son also plays for the same team. Name the other cricketer.

Non Pyramidal Question: Name the cricketer who shares another nickname with a cricketer who is a mentor for the team with the most Indian Premier League titles and whose son plays for the same team.

Category: Sports

Source:

https://en.wikipedia.org/wiki/List_of_nicknames_used_in_cricket

https://www.indiatimes.com/sports/we-all-know-sachin-tendulkar-as-the-master-blaster-but-ever-wondered-how-he-got-the-name-358988.html

Question Answering System	Guess
ChatGPT	No concrete guess
You AI	No concrete guess

Techniques Used: Multi-step, Novel Clues, Common Link

Answer: Vivian Richards

Why is the Question Interesting and difficult?

This question talks about Sir Vivian Richards and Sachin Tendulkar who both are nicknamed 'Master Blaster'. They were both great batsmen of their time. Sachin Tendulkar used to play for the Mumbai Indians, who have won the most number of IPL titles. He is now a mentor to the Mumbai Indians.

This question is difficult because the nickname 'Master Blaster' is a more local (Indian) name for Sachin Tendulkar. Outside India, he is more commonly known as the 'Little Master' or the 'God of Cricket'. On a related note, Rahul Dravid and Steve Waugh are said to have the nickname 'The Wall'. This question is pyramidal which doesn't narrow down on what's being asked until the end of the question.

A person who watches cricket might find this interesting because not all cricketers are given a nickname. In addition to this, Sachin Tendulkar and Sir Vivian Richards are both highly celebrated batsmen and have large fan followings. In India, it would be hard to find a youth who doesn't know Sachin Tendulkar and the same goes for Sir Vivian Richards. Due to Tendulkar's celebrity status in India, his son is also a source of intrigue for many, especially after he was picked by the most successful IPL teams in India. With respect to the popularity of the Indian Premier League, according to Star Sports' post on social media in April, the opening game of IPL 2023 has reached a record number of 140 million people through the live broadcast on the Star Sports Network.

Both ChatGPT and You AI were not able to make a guess. ChatGPT assumes we are talking about Rahul Dravid and Steve Waugh , thus it is unable to use the clue that one of these players' sons plays for the Mumbai Indians. You AI asks for more information, suggesting that it couldn't even list two cricketers with the same nickname.

Question: This player led his country's national team to their first Davis cup title in 2010. His country's capital is the country's financial center. He was born in a city that is also known as the white city and won a bronze medal in 2008 Olympic games and withdrew from the Madrid open recently. Who is this player?

Non-Pyramidal Question: Name the person who won a bronze medal in the 2008 Olympic games for tennis from Serbia.

Category: Sports

Source: https://en.wikipedia.org/wiki/Novak_Djokovic
https://tennis-infinity.com/atp/novak-djokovic-forced-to-pull-out-of-2023-madrid-open#:~:text=The%20Serb%20started%20with%20a,from%20the%202023%20Madrid%20Open.

Question Answering System:	
You	Rafael Nadal
ChatGPT	Rafael Nadal

Answer: Novak Djokovic

Techniques used: Multihop, Semantic Ambiguity, oblique paraphrasing

Why the question is difficult and interesting:

This question is difficult because it uses oblique paraphrasing and does not specify which country's national team this player played for in the Davis cup in 2010. The information about Davis cup would have allowed GPT to infer that the person plays tennis and would have narrowed down the players who play tennis and played for the Davis cup in 2010.

The question uses semantic ambiguity by simply stating that the player was born in a city that is also known as a white city. There are multiple cities across the world that are known as white cities hence the GPT can be confused with other countries, however since it is stated that the national team participated in the davis cup in 2010 and that the player won the bronze medal in 2008 during the olympic games, the GPT should have derived that the national team that the player is playing for is Serbia. Although the question also stated the most up to date information about the player who withdrew from the Madrid Open recently as well, putting together all this information step by step and deriving the answer is tedious, hence making the question difficult and interesting.

This question is correct because it has the factual information quoted from reliable sources. Additionally, this question is pyramidal because it provides a series of hints starting with very broad information such as providing details about when the player's country won the Davis cup title and gradually narrowing down to specific details by stating the specific awards he has received in which games.

Question: This person was born in California and moved to Seattle, and enrolled in university of Washington one year after 1960. He was an actor, filmmaker and he also learnt the traditions of combat for self-defense and military and he grew up in a country that is surrounded by the south china sea.

Non-Pyramidal Question: Name the person who is an American and Hong Kong martial artist.

Category: Sports

Source:
https://brucelee.com/bruce-lee

Question Answering System:	
You	Bruce Lee
ChatGPT	Bruce Lee

Answer: Bruce Lee

Techniques used:
Multihop

Why is the question interesting and difficult:

The question is difficult and interesting because it uses multihop where GPT has to retrieve information about all the people who were born in california and moved to seattle

Question: This woman was the first Indian woman to win a silver medal at the olympics.

and the people who joined university of washington in the year 1961. The sentence is phrased in such a way that it does not directly provide the year that Bruce Lee joined the university which makes it harder and increases the computational time required for GPT to process this question. The question does not directly provide the fact that he is a trained martial artist instead it also uses multihop to introduce the various careers that he had pursued together with martial arts. Additionally it describes the martial arts as traditions of combat for self defense and military instead of directly stating martial arts to induce semantic ambiguity as combats for self defense can be understood as different meanings not necessarily martial arts only . Therefore putting together all this information it is possible for a human to draw these connections and make the correct guess. However, for a GPT, consolidating the information from different sources and matching it to a person in its database is tedious and difficult.

This question is correct because it has the accurate factual information quoted from reliable sources. Additionally, this question is pyramidal because it provides a series of hints starting with very broad information such as providing details about where the actor went for schooling and where he was born and gradually narrowing down to specific details by stating detailed information about the other activities he has been involved in.

She also later went on to win another olympic medal in the next iteration of olympics in the

same sport and category. Where did she win her first olympic medal?

Non Pyramidal Question: Where did PV Sindhu, the first Indian woman to win a silver medal at the olympics, win her first olympic medal?

Category: Sports

Source: https://olympics.com/en/athletes/v-sindhu-pusarla

https://en.wikipedia.org/wiki/P._V._Sindhu

Question Answering System	Guess
ChatGPT	2016 Rio Olympics
Bard	2000 Summer Olympics in Sydney, Australia

Answer: 2016 Rio Olympics

Techniques Used: Multihop, Misdirection

Why is the Question Interesting and difficult?

In this question, we try to ask a question answering system, where PV Sindhu won her first olympic medal in a pyramidal manner. The olympics website and wikipedia both recognize her to be the first Indian woman to win a silver medal at the olympics, making the question factually right. She also won a bronze in the next iteration of the olympics in the same sport and category.

This question is difficult for a computer to answer because firstly because, we use multihop and misdirect the computer into believing we're asking about the person, however we ask about where she had won her first olympic medal. Secondly, the question is pyramidal which doesn't narrow down on what's being asked until the end of the question.

The question also is interesting because we talk about an Indian woman athlete reigning in a sport which is not very popular in the western world, making it more diverse. ChatGPT gets the answer correct but Bard is close and answers Karnam Malleshwari, the first Indian Woman to win a medal at the olympics and gives the wrong answer instead, might be due to the n-gram match with Karnam Malleshwari.

Lifestyle

Question: In a seaside town in the country known to be the home of the largest Mediterranean Island, on Ash Monday, the natives celebrate the first day of Lent by throwing bags of what at friends, family, strangers and visitors?

Non-Pyramidal Question: In Galaxildi, Greece, the natives celebrate the first day of Lent by throwing bags of what at friends, family, strangers and visitors?

Category: Lifestyle

Techniques Used: Multi-step, Novel Clues, Oblique Paraphrasing

Source: https://www.afar.com/magazine/12-crazy-festivals-where-you-can-actually-throw-things-at-people

https://www.washingtonpost.com/lifestyle/kidspost/greeks-let-the-flour-fly-in-clean-monday-tradition/2017/02/28/e151b154-fde0-11e6-8f41-ea6ed597e4ca_story.html

Question Answering System	Guess
ChatGPT	Colored Flour
You AI	Confetti

Answer: Flour

Why this question is interesting and difficult:

This question talks about how Ash Monday is celebrated in a seaside town in Greece. In this question, we attempt to ask a question answering system about what the people of this town throw at each other to celebrate this festival. The facts stated in this question are all available on the film's Wikipedia page.

This question is difficult because it hides the critical clues using novel phrasing. For example, we attempt to hide the name of the country this town is located in, using a novel clue. In addition to this, we change the way of asking questions by asking 'bags of what' (oblique paraphrasing) instead of asking the question in a more formal manner.

You AI is unable to get the answer although it correctly identifies the country that is home to the largest Mediterranean island probably because it needs to be exposed to more information regarding festivals.

Question: This is a food item in one of the most common British breakfasts and is known to be very versatile where it can be eaten raw or cooked. It goes well with bacon and pancakes. Vegans and Jains can't eat this food item.It can be cooked on the outside, with the inside still being uncooked by dropping this raw food item into boiling water that has vinegar in it. This process is called poaching. Another way of cooking this food item is frying it on the pan. What is this food item?

Non-Pyramidal Question: Name the food item that can be cooked in various ways including being poached, scrambled, fried.

Category: Lifestyle

Question Answering System:	
You	Eggs
ChatGPT	Eggs

Answer: Eggs

Techniques used: Common Sense, Common Link, Multi-Step

Why is the question interesting and difficult:

This question is interesting and difficult to answer because it starts off with giving a broad hint that it is a common food item in English breakfasts and could be eaten raw or cooked, which could refer to numerous other food items that fit this description.

The question also uses a common link when it says it goes well with bacon and pancakes, where it is referring to the food item being a good combination with bacon and with pancakes. However since there are many other food items such as french toast, maple syrup, jam or english tea that could be a possible combination with either bacon or pancake, it confuses the system as to what are the possibilities of this food item, hence resulting it to not narrow down to a particular answer.

This question uses common sense because vegans can't eat any products that are from animals and Jains cannot eat this food item as well because of their cultural practice which is common knowledge that everyone is aware of.

The question also uses multihop and gives more information on how this food item can be cooked to introduce GPT to various information related to this product so that it would be more difficult to put together all this information and derive that the final answer is egg.

This question is interesting because it uses a unique way to describe the methods of cooking an egg, and it states the possible combinations that can be taken to consume this food item.

This question is correct because it has the accurate factual information quoted from reliable sources. Additionally, this question is pyramidal because it provides a series of hints starting with very broad information such as providing details about what type of breakfasts this particular food item appears in and gradually narrowing down to specific details by stating the methods of cooking and the other food items that go well with this food item.

Question:

There is a tribe that exists in Asia, where the tribe women adorn rings around their necks. It is to be said that the practice is done to exaggerate sexual dimorphism, and that the physiological disadvantages of the practice outweigh any benefits of this practice. Many argue to stop practicing the adorning of neck rings in hopes that the women can focus on pursuing a higher education to increase their literacy rates. What tribe is this?

Non-pyramidal Question:

There is a tribe that exists in Southeast Asia, where the tribe women adorn rings around their necks. What is this tribe?

Category: Lifestyle

Question Answering System	Guess
https://chat.openai.com/chat	Kayan Lahwi tribe
https://you.com/	Kayan Lahwi tribe

Answer: Kayan Lahwi tribe

Techniques used:
Oblique Paraphrasing, Misleading

Why is this question difficult and interesting?:

The question posed about a unique tribe in Southeast Asia presented a challenge to the AI **Question:** The hostility from these groups of people every time anyone has made contact with them made the government realise that these people want to remain isolated. After the government having tried multiple times to maintain friendly relations, who are the group of people off the coast of India who are being geofenced such that they remain the most isolated group of people in the world?

systems due to the broad nature of the question and the potential difficulty in accurately pinpointing the region. Additionally, the tribe has received less attention in recent years due to political issues in Burma, making it potentially harder to find information about them.

My question is correct as it asks for information about a specific tribe in Southeast Asia while acknowledging the potential difficulty in pinpointing the region and finding information due to the political situation in Burma.

Adversarial writing techniques were not used in this question as the additional context provided about the political situation in Burma was not intended to lead the AI systems astray but rather to acknowledge the potential difficulty in finding information about the tribe.

The question answering systems likely struggled to provide an accurate answer due to the lack of information available about the tribe and the potential difficulty in accurately pinpointing the region in Southeast Asia.

The question is pyramidal to some extent as it starts with a general statement about the lack of attention towards the tribe in recent years before narrowing down to the specific challenge of pinpointing the region in Southeast Asia.

Non Pyramidal Question: Which group of people off the coast of India has the Indian government geofenced to keep them the most isolated people in the world?

Category: Lifestyle

Source:
https://en.wikipedia.org/wiki/Sentinelese

https://en.wikipedia.org/wiki/Uncontacted_peoples

Question Answering System	Guess
ChatGPT	Sentinelese
Bard	Sentinelese

Answer: Sentinelese

Techniques Used: Paraphrasing, Multihop, Misdirection

Why is the question interesting and difficult?

This question is interesting because it talks about a group of people still having no human contact with the outside world in this age of

technology and digital communication. This question talks about an islander tribe in the Andaman and Nicobar Islands in India. The sentinelese of the North Sentinel Island have not maintained contact with the outside world for 1000's of years now and want it to remain that way.

The government of India has made sure that they are geofenced and have their interests protected as they wish, and to preserve their community from the outside world of diseases they haven't been exposed to. This question employs paraphrasing, multihop and misdirection by talking about the tribe being contacted and yet remaining the most isolated and uncontacted people in the world.

Both the answering systems get it right because there is only one tribe in the world that doesn't want any humans to contact them and reserve their spot as the most isolated people.

Question: This disease was found in a 2004 study to have increased prevalence in India due to changes in lifestyle and environment brought about by Industrialisation and urbanization? What disease is treated by a drug that was recently approved to treat obesity and this disease?

Non Pyramidal Question: What disease is treated by semaglutide other than obesity?

Category: Lifestyle

Source:

https://en.wikipedia.org/wiki/Diabetes_in_India

https://en.wikipedia.org/wiki/Semaglutide

https://en.wikipedia.org/wiki/Diabetes

Question Answering System	Guess

ChatGPT	Diabetes
Bard	Type 2 Diabetes

Answer: Type 2 Diabetes

Techniques Used: Novel Clues, Multihop

Why is the question interesting and difficult?

The question is an interesting question because it talks about the lifestyle disease of type 2 diabetes and the reasons why it occurs in Indians as a consequence of changed lifestyle post Industrialisation and urbanization. A drug called WeGovy or Semaglutide was recently approved by the US FDA to treat type 2 diabetes and was also found to help with weight loss, hence being approved for use as anti obesity treatment as well.

We make use these facts in a multihop fashion to describe type 2 diabetes being found to be a disease occuring in higher prevalence due to a change observed in lifestyle in Indians after urbanisation and industrialisation. We use the novel clue of drug that was approved for treating this disease and obesity too.

This is possibly why ChatGPT comes close but doesn't answer the question right. GPT doesn't seem to be performing extremely well with novel clues, however, Bard seems to get it right because its trained in real time on the latest information.

Music

Question: This person was born in Pennsylvania, and started her career from the country that is home to Vanderbilt University. She released her second studio album in 2008 and her third album in 2010 which was produced by an American independent record label. She has won 12 grammy awards and 40 american music awards till current date. Who is this person?

Non-Pyramidal Question: Name the singer who is known for her popular songs Blank Space and Shake it off.

Category: Music

Source:
https://en.wikipedia.org/wiki/Taylor_Swift

Question Answering System:	
You	Taylor Swift
ChatGPT	Taylor Swift

Answer: Taylor Swift

Techniques Used: Multihop, Oblique Paraphrasing

Why is the question interesting and difficult:

This question is interesting and difficult because the question uses oblique paraphrasing by simply stating the location where the person was born and where she started her career only. There could be multiple people who fit this description, hence it is still unclear to make a decision about who this person is. The question also uses multihop to include more information

about the person, by stating the dates of her album and who it was produced by, hence although it gives GPT more information about the person, GPT has to retrieve information from multiple sources and consolidate all these details and match it to a person. Additionally since there might be few people who won 12 grammy awards and 40 american music awards, GPT should consolidate all the information that it has received in the question and narrow down the possibilities and derive the answer which is more difficult.

This question is correct because it has the accurate factual information quoted from reliable sources. Additionally, this question is pyramidal because it provides a series of hints starting with very broad information such as providing details about the location the singer was born and gradually narrowing down to specific details by stating the number of awards she received the specific albums she has done.

Question:
What is the singing style that utilises 1 or 2 pitches simultaneously, that is done in regions of Siberia, Tuva and that the style represents ancient pastoral animism, as well as being recognised in the Representative List of the Intangible Cultural Heritage of Humanity of UNESCO?

Non-pyramidal Question: What is the type of singing that utilizes 1 or 2 pitches simultaneously that is done in regions of Tuva and Siberia?

Category: Music

Source:
https://en.wikipedia.org/wiki/Tuvan_thro at_singing
https://www.alashensemble.com/about_tts .htm

Question Answering System	Guess

https://chat.openai.com/chat	Tuvan Throat Singing
https://you.com/	Tuvan Throat Singing

Answer: Tuvan Throat Singing

Techniques used: Oblique Paraphrasing, Misleading

Why is this question difficult and interesting?:
The question posed to the AI systems about singing techniques in the Central Asian region presented a challenge due to the wide variety of techniques used in this region. By adding Serbia to the question, there were additional possibilities for music and singing techniques, which could potentially have confused the AI systems. However, the systems were able to accurately identify the correct singing techniques despite this potential for confusion.

Question:
There is a song that originates from Cambodia, where it is sung in the Surin dialect of the languages spoken in Cambodia (Khmer). This song was found in a random cassette tape found in the US around the 2000s, and attempts of translating it into English and Czech were seen on youtube. The song talks about frogs and lizards stealing marijuana, as well as hares knowing only how to drink. What is the name of this song?

Non-Pyramidal Question:

My question is correct as it asks for specific information about singing techniques in the Central Asian region while acknowledging the potential for confusion with the inclusion of Serbia in the question.

Adversarial writing techniques were not used in this question as the additional context provided by Serbia was not intended to lead the AI systems astray but rather to acknowledge the potential for confusion.

The question answering systems performed well in this case, likely due to their ability to identify and categorize different types of singing techniques based on the provided context.

The question is not pyramidal as it does not gradually add more specific details or context to narrow down the search space. However, the acknowledgement of potential confusion with the inclusion of Serbia in the question does demonstrate a level of awareness and consideration for the limitations of the AI systems.

There is a song that originates from Cambodia, where it is sung in the Surin dialect of the languages spoken in Cambodia (Khmer). The song talks about frogs and lizards stealing marijuana, as well as hares knowing only how to drink. What is the name of this song?

Category: Music

Source:
https://lyricstranslate.com/en/sat-tee-touy-look-owl-sat-tee-touy.html

Question Answering System:	
You	**Couldn't answer**
ChatGPT	**Couldn't answer**

Answer: Sat Tee Touy

Techniques Used: Oblique Phrasing, Misleading

Why is the question interesting and difficult:

This question is difficult as the song is obscure, and it is obscured with different languages such as Khmer, Surin, English, and Czech, which makes it harder for the AI systems to figure out the title of the music. Even though this song is obscure, it should be easy to find on the internet due to its esoteric nature of the lyrics. However, the AI systems are not able to find the song's name, indicating that this question may be more challenging than anticipated. The adversarial writing techniques used in this question include the intentional use of obscure language, as well as the assumption that the esoteric nature of the lyrics would make it easier to find the song. Nonetheless, the AI systems' inability to provide an answer suggests that the question may be even more difficult than originally thought.

Question: This song went on to win the Oscar for Best Original Song in 2023. Who is the composer of the golden globe award-winning song that was performed live at the Oscars in 2023?

Non Pyramidal Question: Who is the composer of the song 'Naatu Naatu'?

Category: Music

Source:
https://en.wikipedia.org/wiki/Naatu_Naatu

Question Answering System	Guess
ChatGPT	No Answer
Bard	M.M. Keeravani

Answer: M.M. Keeravani

Techniques Used: Multihop, Novel Clues

Why is this question interesting and difficult?

The question is interesting because it talks about the Telugu Film Industry and its music, which has only gained popularity in the recent years for its filmmaking and craft through the recognition in the Oscars and the Golden Globe Awards.

A knowledgeable human can easily tell that the questions probe for the composer of the song rather than the song itself. Any ordinary Indian Movie watcher can easily guess the song but would have to google the composer's names because its something only the native listeners would know.

The question employs multihop and asks about recent events making it difficult for a computer to answer. Moreover, the question is also pyramidal as it begins with a broad base of probing about a song that won an oscar, then narrows down to the composer of the song performed at the event and asks for their name.

ChatGPT gets the question wrong due to use of novel clues, however Bard is being trained on new clues on a daily basis, helping it get the question right.

Question:

The director for the film 'comali' directed a hit movie that was released last year in multiple languages. This director also acted in this hit movie. The music album for this movie became very popular because of the groovy beats. Who is the music composer of this movie?

Non-Pyramidal Question: Name the music composer who composed the Rowdy Baby song from Maari 2.

Category: Music

Source:

Question Answering System:	
You	Sean Roldan
ChatGPT	Hiphop Tamizha

Answer: Yuvan Shanker Raja

Techniques Used: Common link and oblique paraphrasing

Why is the question interesting and difficult:

This question is difficult because it uses a common link to say that the director for the movie stated in the question is the same director for the movie asked in the question, and it is specified that this movie was released last year. However GPT interprets it as 2019 and returns the music composer of another movie called Mehendi circus directed by the same director. Additionally because of the oblique paraphrasing of the question, which is stating that the music album became really popular and that the director also acted in this movie does not explicitly state which movie the question is referring to and the role of the director when he acted in the movie. Therefore, understanding the question's details and narrowing down the possible music composers with the given information becomes more difficult.

This question is correct because it has the accurate factual information quoted from reliable sources. Additionally, this question is pyramidal because it provides a series of hints starting with very broad information such as providing details about the movie that he composed the music and gradually narrowing down to specific details.