

Conditional Generation with a Question-Answering Blueprint

Mirella Lapata

School of Informatics
University of Edinburgh



Conditional Generation: The Wish List

- Generate natural language towards a **communicative goal**
- which is **faithful** and can be **attributed** to its sources
- while users explicitly **control** generation outcome
- without being sensitive to the **modality** of the input

Conditional Generation: Long-form QA



The screenshot displays two search results side-by-side. On the left, the NOAA Office of History website lists several facts about the Titanic, such as its maiden voyage, sinking, and cultural impact. On the right, the Encyclopaedia Britannica article 'Why is the Titanic famous?' provides a detailed narrative about the ship's fame and its sinking.

- Some queries have only long-form answers
- based on input from multiple documents
- where hallucinations and attribution can be more problematic

Conditional Generation: Dialogue Summarization

The Justice League Recombination, The Big Bang Theory, S4.E11



Penny starts dating Zack again. Used to being the target of bullying, Leonard, Sheldon, Howard and Raj instead bully Zack through their implications to him that he's stupid. Not fully realizing what they're doing until Penny points it out to them, the guys decide to apologize to Zack in the only way they can. Accepting their apology, Zack decides to start hanging out with them, much to Penny's chagrin. This hanging out leads to the idea that Zack should replace Leonard as Superman in their group Justice League of America costume for the New Year's Eve party at the comic book store, with Penny the sixth wheel as Wonder Woman. These actions lead to questions from Leonard to Penny about why she's back together with Zack.

Conditional Generation: Visual Storytelling

Image Sequence:



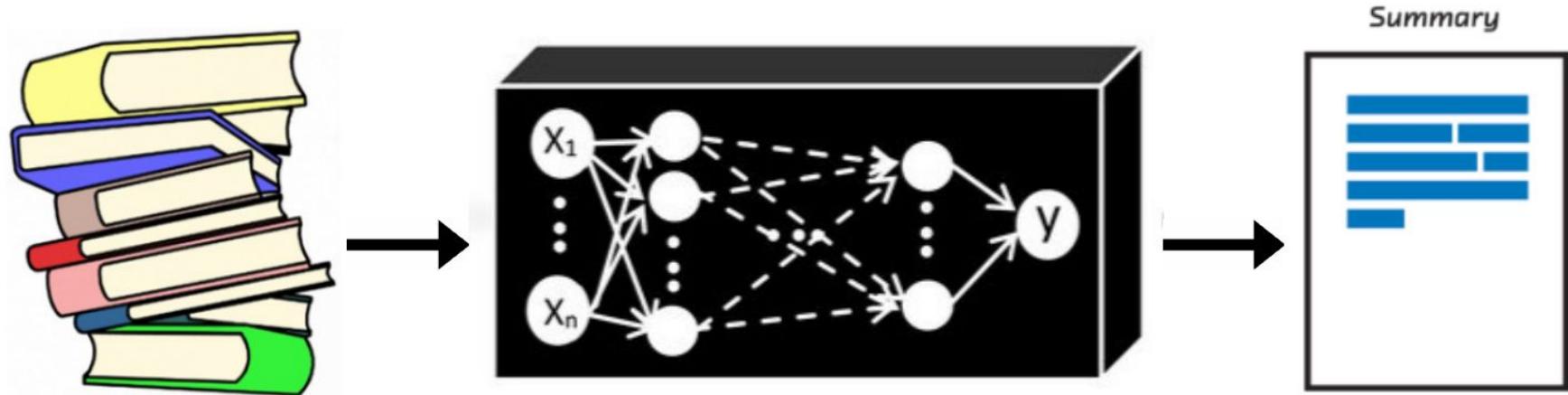
Gold-standard Story:

It was customary to give big brass a full tour. Of course, the mess hall would usually serve something a little better than average on "Big Brass" days. He looked like he was okay with it. After lunch the tour continued. He was introduced to the Commanding officers.

Story Generated by KE-VIST:

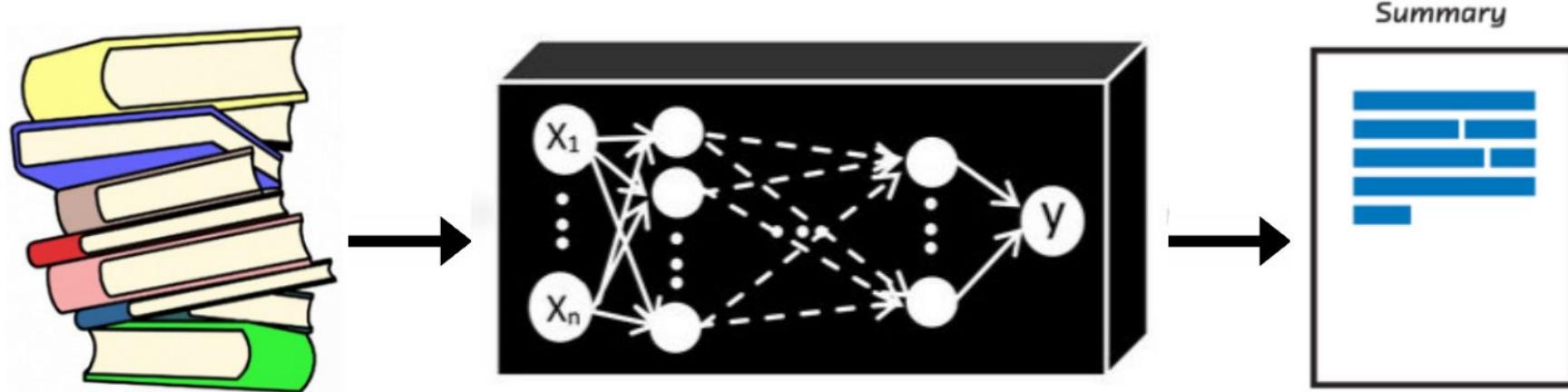
We all met up with the men. They walked a lot in the man walking. We had a lot of food. I gave a

The Blackbox Nature of Conditional Generation



- Change the way **entities** are represented (Puduppully et al., 2019; Iso et al., 2019)
- The **decoder** skips low-confidence tokens (Tian et al., 2019)
- **Encode** documents hierarchically (Rhode et al., 2021)
- Adopt **sparse attention mechanisms** (Child et al., 2019; Beltagy et al., 2020)

The Blackbox Nature of Conditional Generation



Introduce intermediate **planning** components
(Wiseman et al., 2018; Moryossef et al., 2019;
Puduppully et al., 2022; Narayan et al., 2022)

The QUD Model of Discourse

Questions under Discussion (QUD) theory of discourse structure (Carlson, 1983; Ginzburg, 1994; Van Kuppevelt, 1995; Larson, 2002; Roberts, 2012; Riester, 2019).

A **QUD** is a partially structured set of questions which discourse participants are mutually committed to resolving at a given point in time.

A **question** is a set of alternative possibilities (corresponding in some way to the possible answers to the question).

Discourse contains **implicit questions** for each of the assertions made which are turned into answers.

Question-Answering Blueprints as Content Plans



what is the titanic known for



What kind of ship was the RMS Titanic? British passenger liner

What line operated the RMS Titanic? the White Star Line

How many ships did the White Star Line make? three ships

In what ocean did the RMS Titanic sink? Pacific Ocean

When did the RMS Titanic hit an iceberg? On its maiden voyage

The screenshot shows a Wikipedia page for the RMS Titanic. The sidebar on the left contains a list of related topics such as Biological Diversity, Carbon Capture Storage, Data Buoy Vane, Fisheries, Geoengineering, Heritage, IMO Area Based Management Tools, and Jurisdiction. Below this is a list of 8 related questions:

1. Biological Diver...
2. [Titanic](#)
3. 'Unsinkable' Fatal Flaw
4. Passengers of the Titanic
5. Titanic Sinking
6. The Titanic and the Iceberg
7. Titanic's last message
8. [Wikivoyage](#)

The main content area discusses the sinking of the Titanic and its enduring appeal. It mentions that interest in the ship increased after the release of the 1997 film "Titanic".

Related Q&A

 Why did the <i>Titanic</i> sink? TOP QUESTION	 How many people died when the <i>Titanic</i> sank? TOP QUESTION
 Where is the wreck of the <i>Titanic</i> ? TOP QUESTION	

Question-Answering Blueprints as Content Plans



what is the titanic known for



What kind of ship was the RMS Titanic? British passenger liner

What line operated the RMS Titanic? the White Star Line

How many ships did the White Star Line make? three ships

In what ocean did the RMS Titanic sink? Pacific Ocean

When did the RMS Titanic hit an iceberg? On its maiden voyage

The screenshot shows the Wikipedia page for the RMS Titanic. The page title is "RMS Titanic". The main content discusses the ship's construction, its maiden voyage, and its sinking. Below the main text are sections for "Top Questions", "Facts & Related Content", "Media", and "More". To the right, there is a sidebar with related questions like "Why did the Titanic sink?", "How many people died when the Titanic sank?", and "Where is the wreck of the Titanic?". The bottom of the page includes a "Print/export" section and links to other projects like Wikimedia Commons and Wikinews.

Biological Diversity
Carbon Capture Storage
Data Buoy Vane
Fisheries
Geoengineering
Heritage
IMO Area Based Management Tools
Jurisdiction

3. 'Unsinkable' Fatal Flaw
4. Passengers on the Titanic
5. Titanic Sank
6. The Titanic and the Iceberg
7. Titanic's Last Voyage
8. Titanic Sank

TOP QUESTIONS

Facts & Related Content

Quizzes

Media

Videos

Images

More

Print/export

Download as PDF

Printable version

In other projects

Wikimedia Commons

Wikinews

Wikiquote

* Wikivoyage

What links here

Related channels

Special pages

Permanent link

Page information

Cite this page

Wikidata item

of its wreckage, interest in the *Titanic* only increased, and its enduring appeal was evident with the huge success of James Cameron's 1997 film about the doomed ocean liner.

Related Q&A

Why did the *Titanic* sink?
TOP QUESTION

How many people died when the *Titanic* sank?
TOP QUESTION

Where is the wreck of the *Titanic*?
TOP QUESTION

Question-Answering Blueprints as Content Plans



what is the titanic known for



What kind of ship was the RMS Titanic? British passenger liner

What line operated the RMS Titanic? the White Star Line

The screenshot shows a search interface with the query "what is the titanic known for". Below the search bar, there are several search results. One result is highlighted with a red box and the text "Why is the Titanic famous?". Other visible results include "The Build", "Unsinkable", "Fatal Flaw", "Passenger", "Titanic", "Sink", "The Titan", "Iceberg", "Titanic's", and "Titanic S". The results are presented in a grid-like layout with various icons and text snippets.

In what ocean did the RMS Titanic sink? Pacific Ocean

When did the RMS Titanic hit an iceberg? On its maiden voyage

The screenshot shows the Wikipedia article for the RMS Titanic. The page includes a sidebar with "Areas of Practice" (Arctic, Antarctic, etc.) and a main content area with sections like "Recent changes", "Key Terms", "Top Questions" (listing "Why is the Titanic famous?", "Facts & Related Content", and "Quizzes"), "Media" (Videos, Images), and "More" (Wikimedia Commons, Wikinews, Wikiquote, Wikivoyage). Below the main content, there are "Related Q&A" boxes for questions like "Why did the Titanic sink?", "How many people died when the Titanic sank?", and "Where is the wreck of the Titanic?". The page also features decorative floral illustrations.

Question-Answering Blueprints as Content Plans



what is the titanic known for



What kind of ship was the RMS Titanic? British passenger liner

What line operated the RMS Titanic? the White Star Line

The screenshot shows a search interface with the query "what is the titanic known for". The top result is a link to the Wikipedia article on the RMS Titanic, titled "Why is the *Titanic* famous?". The page includes navigation links like "Offices", "Links", and "WIKIPEDIA The Free Encyclopedia".

In what ocean did the RMS Titanic sink? N. Atlantic Ocean

When did the RMS Titanic hit an iceberg? On its maiden voyage

The screenshot shows the full Wikipedia article for the RMS Titanic. The main content discusses the sinking of the ship, mentioning it was on its maiden voyage when it hit an iceberg in the North Atlantic. The sidebar contains sections like "Areas of Practice", "Recent changes", "Key Terms", "Comprehension Quiz", "Top Questions", "Facts & Related Content", "Quizzes", "Media", "Videos", "Images", and "More". Below the main content, there are "Related Q&A" sections with questions about the sinking and the number of deaths.

Question-Answering Blueprints as Content Plans



what is the titanic known for



What kind of ship was the RMS Titanic? British passenger liner

What line operated the RMS Titanic? the White Star Line

In what ocean did the RMS Titanic sink? N. Atlantic Ocean

When did the RMS Titanic hit an iceberg? On its maiden voyage

Fast Facts

Comprehension Quiz

Top Questions

Facts & Related Content

Quizzes

Media

Videos

Images

More

More Articles On This Topic

Researcher's Note

Additional Reading

Contributors

Article History

Why did the *Titanic* sink?
TOP QUESTION

Where is the wreck of the *Titanic*?
TOP QUESTION

How many people died when the *Titanic* sank?
TOP QUESTION

Question-Answering Blueprints as Content Plans



what is the titanic known for



What kind of ship was the RMS Titanic? British passenger liner

What line operated the RMS Titanic? the White Star Line

In what ocean did the RMS Titanic sink? N. Atlantic Ocean

When did the RMS Titanic hit an iceberg? On its maiden voyage

When did the DiCaprio movie come out? 1997

The screenshot shows a Wikipedia page for the RMS Titanic. The page title is "RMS Titanic". The main content discusses the ship's construction, its voyage, and its sinking. A sidebar on the left provides links to related topics like Biological Diversity, Carbon Capture Storage, and Geoengineering. Another sidebar on the right lists "Top Questions" such as "Why did the Titanic sink?", "How many people died when the Titanic sank?", and "Where is the wreck of the Titanic?". The page footer includes standard Wikipedia navigation links.

Question-Answering Blueprints as Content Plans



what is the titanic known for



What kind of ship was the RMS Titanic? British passenger liner

What line operated the RMS Titanic? the White Star Line

In what ocean did the RMS Titanic sink? N. Atlantic Ocean

When did the RMS Titanic hit an iceberg? On its maiden voyage

When did the DiCaprio movie come out? 1997

↳ Binational Divers ↳ THE BRIEF What links here > Top Questions

of its wreckage, interest in the *Titanic* only increased, and its enduring appeal was evident

RMS Titanic was a British passenger liner, operated by the White Star Line. It sank in the North Atlantic Ocean after striking an iceberg during its maiden voyage. The 1997 film *Titanic* starring Leonardo Di Caprio has popularized the disaster and its aftermath.

Question-Answering Blueprints as Content Plans

- Blueprints as intermediate (discrete) planning stage for conditional generation
- Reduce **faithfulness** errors and increase **controllability**
- Are better for **long-form** inputs and outputs

Where do the blueprints come from?

Datasets as we know them consist of document(s)-output pairs!

Training Data Annotation with Blueprints

Shelby Mustang

⋮ A 19 languages ▾

Read Edit View history Tools ▾

From Wikipedia, the free encyclopedia

The **Shelby Mustang** is a high-performance variant of the [Ford Mustang](#) built by [Shelby American](#) from 1965 to 1967 and by the [Ford Motor Company](#) from 1968 to 1970.

In 2005, Ford revived the Shelby nameplate for a high-performance model of the fifth-generation Ford Mustang.^[1]

1st generation (1965–1970) [edit]

1965–1966 Shelby GT350 [edit]

The 1965–1966 cars were the smallest and lightest of the GT 350 models. Production started when 289 AC Cobra production was wrapping up and 427 AC Cobra production was beginning, both of which were [Ford-powered AC-based two-seat sports car](#) also produced by [Shelby American](#) during the same period. Both models use the Cobra emblem, similar paint scheme, and the optional "Cobra" valve covers on many GT350s that were part of a marketing tie-in by Shelby, as well as one of his iconic symbols. All 1965–66 cars had the [Windsor 289 cu in \(4.7 L\) HiPo K-Code 271 hp \(202 kW; 275 PS\) V8 engine](#), modified with a large 4-barrel [Holley 715 CFM carburetor](#) to produce 306 bhp (310 PS; 228 kW) at 6,000 rpm and

Shelby Mustang



1968 Shelby Cobra GT350

Overview

Manufacturer Ford Performance
Shelby
SVT

Also called Ford Mustang Shelby
Ford Shelby Mustang
Mustang Shelby
Mustang GT500

The Shelby Mustang is a high performance variant of the Ford Mustang which was built by Shelby American from 1965 to 1968, and from 1969 to 1970 by Ford. Following the introduction of the fifth generation Ford Mustang in 2005, the Shelby nameplate was revived as a new high-performance model, this time designed and built by Ford.

Training Data Annotation with Blueprints

The [Shelby Mustang](#) is a high performance variant of the [Ford Mustang](#) which was built by [Shelby American](#) from [1965](#) to [1968](#), and from [1969](#) to [1970](#) by Ford. Following the introduction of the fifth generation [Ford Mustang](#) in [2005](#), the [Shelby nameplate](#) was revived as a new high-performance [model](#), this time designed and built by [Ford](#).

Question-Answering Blueprints as Content Plans

The **Shelby Mustang** is a high performance variant of the **Ford Mustang** which was built by **Shelby American** from **1965** to **1968**, and from **1969** to **1970** by Ford. Following the introduction of the fifth generation **Ford Mustang** in **2005**, the **Shelby nameplate** was revived as a new high-performance **model**, this time designed and built by **Ford**.

What is a high performance variant of the Ford Mustang?	Shelby
Who built the Shelby Mustang from 1969 to 1970?	Ford
Q2: What is the high performance variant of the Ford Mustang called?	Shelby
What is a high performance variant of the Ford Mustang?	Shelby Mustang
What is a high performance variant of the Ford Mustang?	Shelby
What is a Shelby Mustang?	A high performance variant
The Shelby Mustang is a high performance variant of what?	Ford Mustang
During what years was the Shelby Mustang built by Shelby American?	1965 to 1968
Who built the Shelby Mustang from 1965 to 1968?	Shelby American
In what year was the fifth generation of the for Mustang introduced?	2005
What generation of Mustang was introduced in 2005?	the fifth generation
What was the Shelby Mustang revived as?	a new high-performance model

Question-Answering Blueprints as Content Plans

The **Shelby Mustang** is a high performance variant of the **Ford Mustang** which was built by **Shelby American** from **1965** to **1968**, and from **1969** to **1970** by **Ford**. Following the introduction of the fifth generation **Ford Mustang** in **2005**, the **Shelby** nameplate was revived as a new high-performance **model**, this time designed and built by **Ford**.

What is a high performance variant of the Ford Mustang?	Shelby
Who built the Shelby Mustang from 1969 to 1970?	Ford
Q2: What is the high performance variant of the Ford Mustang called?	Shelby
What is a high performance variant of the Ford Mustang?	Shelby Mustang
What is a high performance variant of the Ford Mustang?	Shelby
What is a Shelby Mustang?	A high performance variant
The Shelby Mustang is a high performance variant of what?	Ford Mustang
During what years was the Shelby Mustang built by Shelby American?	1965 to 1968
Who built the Shelby Mustang from 1965 to 1968?	Shelby American
In what year was the fifth generation of the for Mustang introduced?	2005
What generation of Mustang was introduced in 2005?	the fifth generation
What was the Shelby Mustang revived as?	a new high-performance model

Question-Answering Blueprints as Content Plans

The **Shelby Mustang** is a high performance variant of the **Ford Mustang** which was built by **Shelby American** from **1965** to **1968**, and from **1969** to **1970** by Ford. Following the introduction of the fifth generation **Ford Mustang** in **2005**, the **Shelby** nameplate was revived as a new high-performance **model**, this time designed and built by **Ford**.

Who built the Shelby Mustang from 1969 to 1970? Ford

During what years was the Shelby Mustang built by Shelby American? 1965 to 1968

In what year was the fifth generation of the for Mustang introduced? 2005

What was the Shelby Mustang revived as? a new high-performance model

Question-Answering Blueprints as Content Plans

Shelby Mustang

⋮ A 19 languages ▾

Article Talk

Read Edit View history Tools ▾

From Wikipedia, the free encyclopedia

The **Shelby Mustang** is a high-performance variant of the [Ford Mustang](#) built by [Shelby American](#) from 1965 to 1967 and by the [Ford Motor Company](#) from 1968 to 1970.

In 2005, Ford revived the Shelby nameplate for a high-performance model of the fifth-generation Ford Mustang.^[1]



1st generation (1965–1970) [edit]

1965–1966 Shelby GT350 [edit]

The 1965–1966 cars were the smallest and lightest of the GT 350 models.

Production started when 289 AC Cobra production was wrapping up and 427 AC Cobra production was beginning, both of which were [Ford](#)-powered AC-based two-seat sports car also produced by [Shelby American](#) during the same period. Both models use the Cobra emblem, similar paint scheme, and the optional "Cobra" valv covers on many GT350s that were part of a marketing tie-in by Shelby, as well as one of his iconic symbols. All 1965–66 cars had the [Windsor 289 cu in \(4.7 L\) HiPo K-Code 271 hp \(202 kW; 275 PS\) V8 engine](#), modified with a large 4-barrel Holley

Who built the Shelby Mustang from 1969 to 1970? Ford

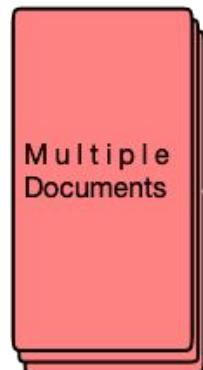
During what years was the Shelby Mustang built by Shelby American? 1965 to 1968

In what year was the fifth generation of the for Mustang introduced? 2005

What was the Shelby Mustang revived as? a new high-performance model

The Shelby Mustang is a high performance variant of the Ford Mustang which was built by Shelby American from 1965 to 1968, and from 1969 to 1970 by Ford. Following the introduction of the fifth generation Ford Mustang in 2005, the Shelby nameplate was revived as a new high-performance model, this time designed and built by Ford.

End-to-End Blueprint Model



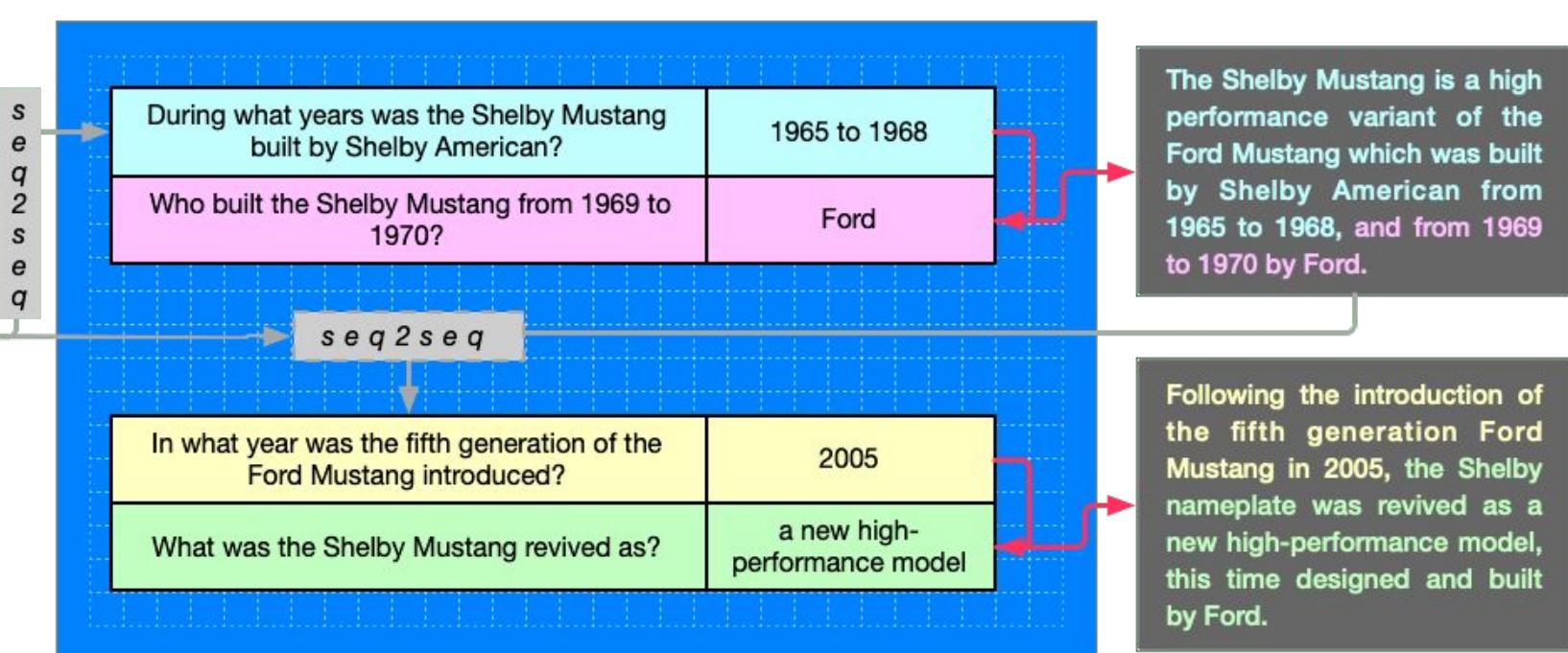
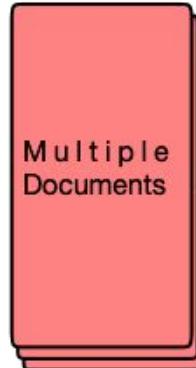
seq
2
seq

During what years was the Shelby Mustang built by Shelby American?	1965 to 1968
Who built the Shelby Mustang from 1969 to 1970?	Ford
In what year was the fifth generation of the Ford Mustang introduced?	2005
What was the Shelby Mustang revived as?	a new high-performance model

The Shelby Mustang is a high performance variant of the Ford Mustang which was built by Shelby American from 1965 to 1968, and from 1969 to 1970 by Ford. Following the introduction of the fifth generation Ford Mustang in 2005, the Shelby nameplate was revived as a new high-performance model, this time designed and built by Ford.

Generates blueprint and output sequence in one go.

Iterative Blueprint Model



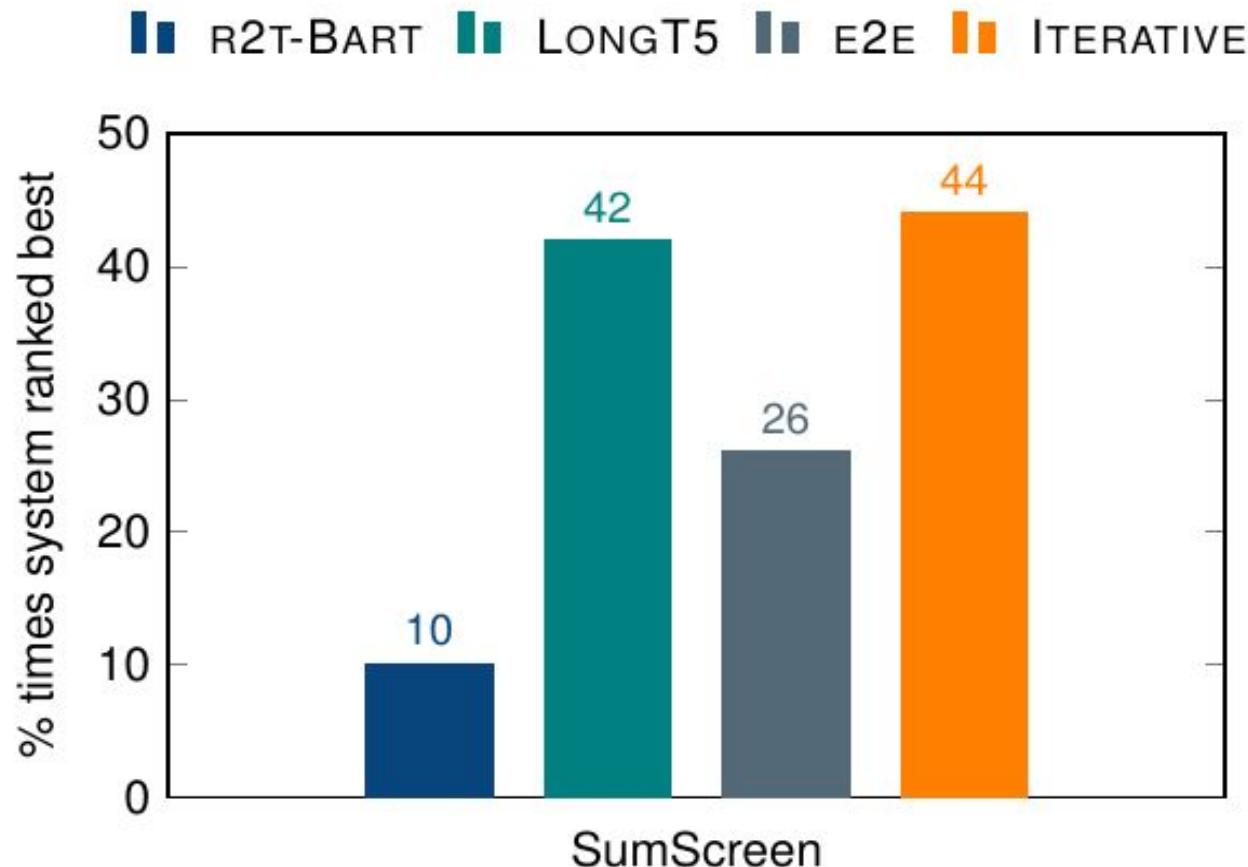
Iteratively plans and generates one sentence at a time.

Experimental Setting

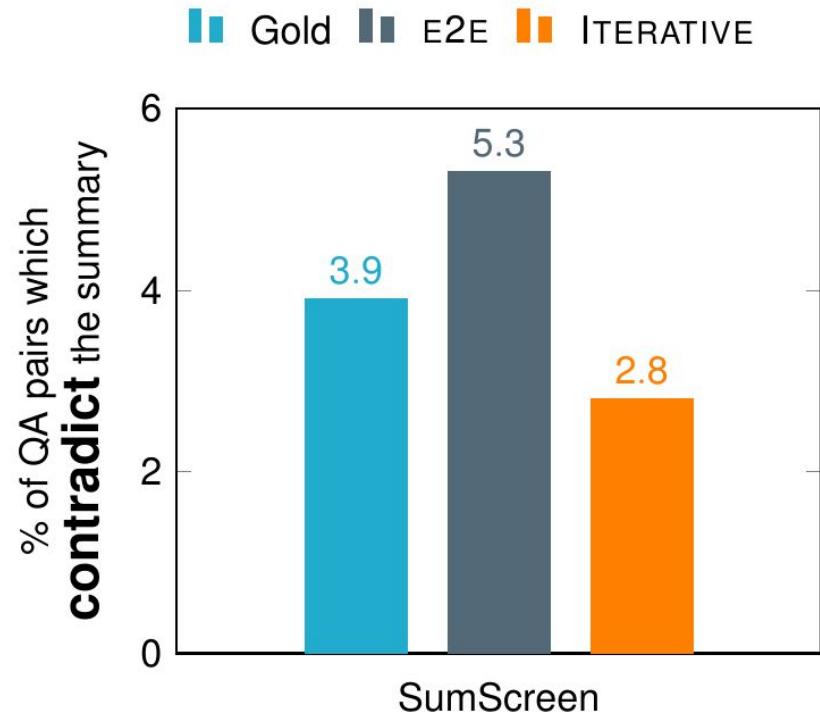
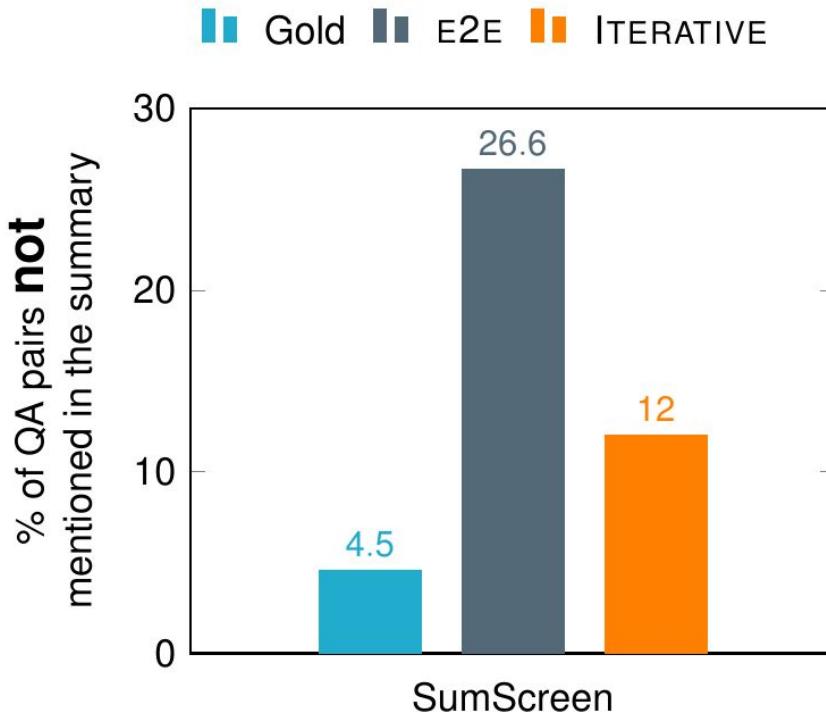
- **AQuaMuse** (Kulkarni et al., 2002, 2021): long-form question answering, simulates search engine, answer based on multiple retrieved documents.
- **WikiCatSum** (Perez-Beltrachini et al., 2019): topic-focused multi-document summarization, generate Wikipedia abstracts.
- **SummScreen** (Chen et al., 2022): dialogue summarization, generate summaries of TV episodes (e.g., CSI, The Bing Bang Theory).

All experiments with a LONGT5 (Guo et al., 2021) backbone model.

Evaluation: How Good are the Blueprint Summaries?



Evaluation: Are Summaries Grounded to their Blueprint?



Interim Conclusions

- Blueprint models are more informative than vanilla seq2seq models
- Errors can be examined and traced back to the blueprint
- QA pairs are intuitive and user-friendly, human-in-the-loop generation
- How do blueprints work across tasks and modalities?

Visual Story Telling: Blueprint Annotation



My Uncle Jack made us go to the Purdue game on Saturday. He went to Purdue and he thinks they are great no matter what else anyone thinks. We all had to sit in the bleachers and wear some ridiculous brown coats...

Q₀: Where did my uncle Jack made us go on Saturday?

A₀: the Purdue game

Q₁: When did my uncle Jack made us go to the Purdue game?

A₁: Saturday

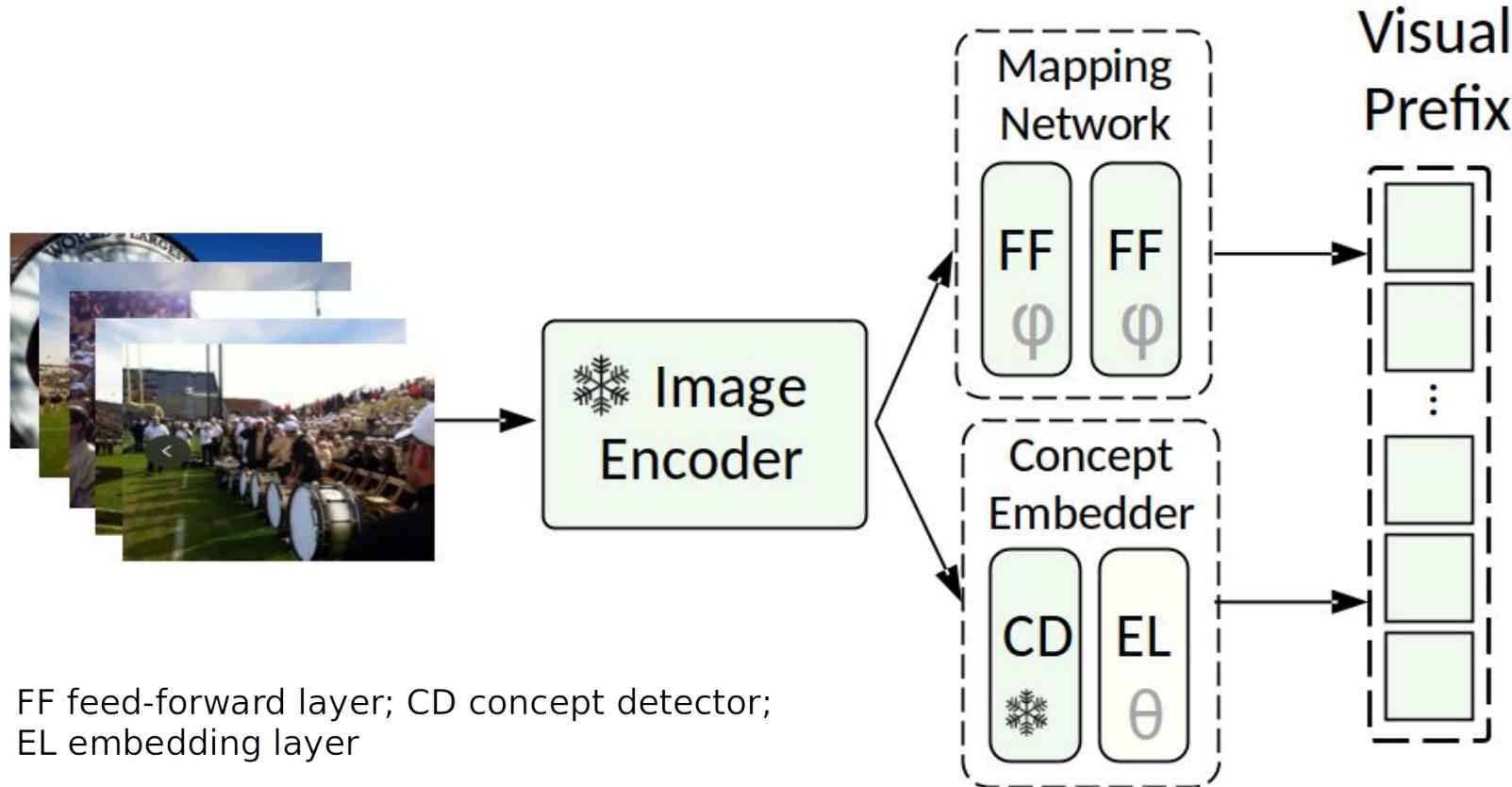
Q₂: Where did we all sit?

A₂: the bleachers

Q₃: What did the fans of Purdue have to wear?

A₃: some ridiculous brown coats

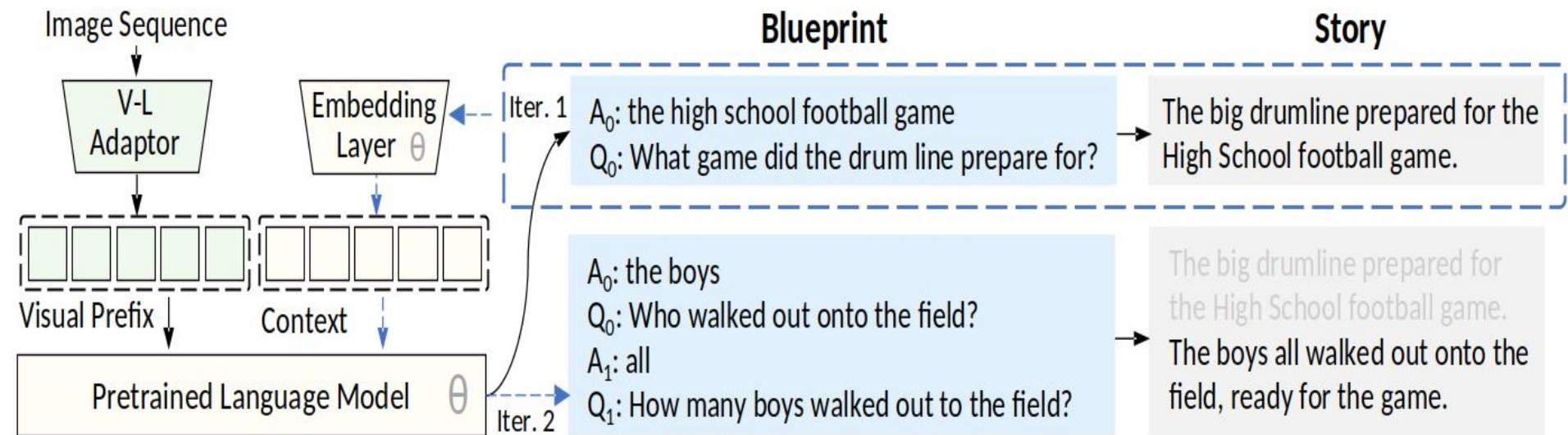
Visual Prefixes



Blueprint Storytelling Model (Iterative)

- Use a generation model (based on T5) to turn VIST stories into blueprints
- Train a linear model to map image sequences into visual prefixes (based on ResNet-152)
- Fine-tune a pretrained model to generate blueprints and stories from visual prefixes
- Generation proceeds sentence-by-sentence, where previous sentences are context for the current one: iterative model
- An end-to-end model that generates all sentences at once from the visual prefix performs worse

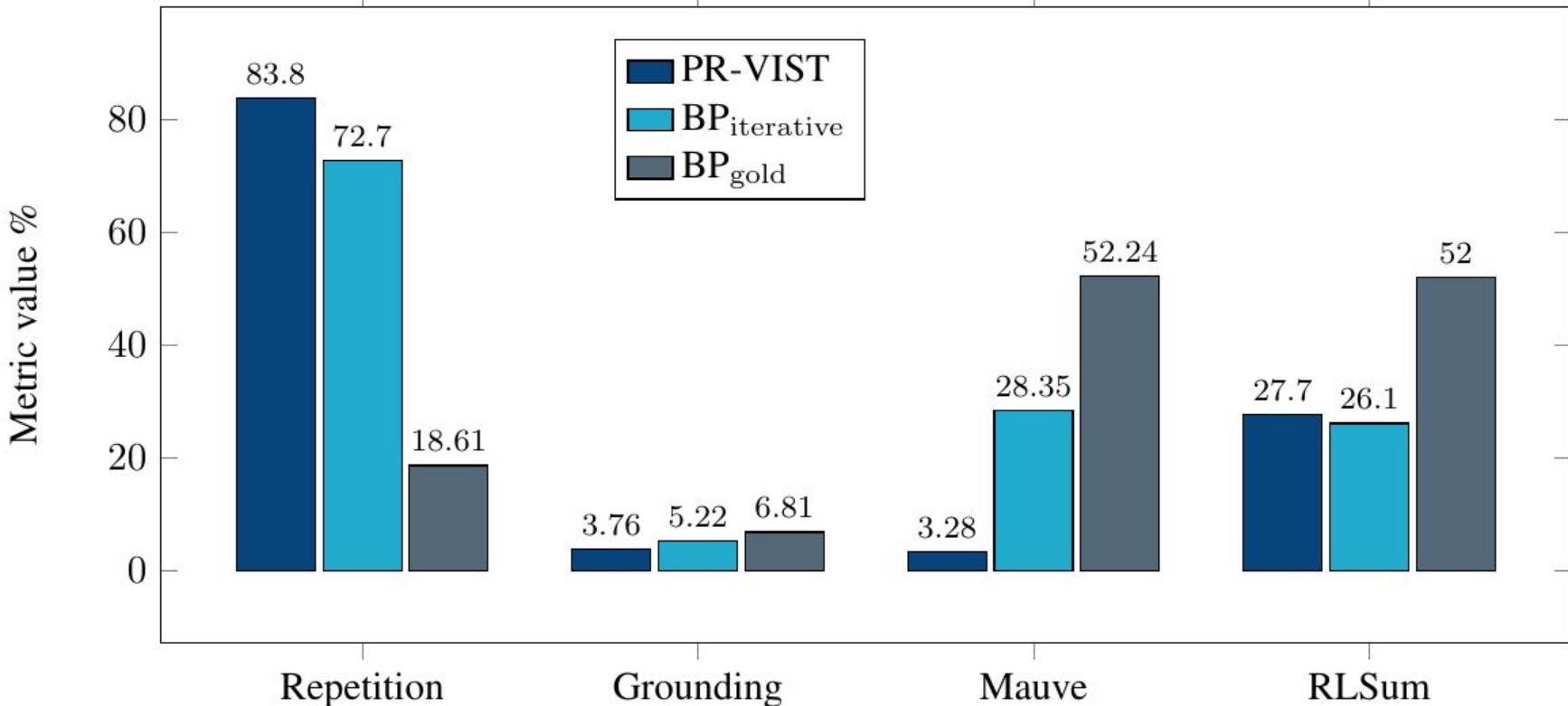
Iterative Blueprint Model



Evaluation

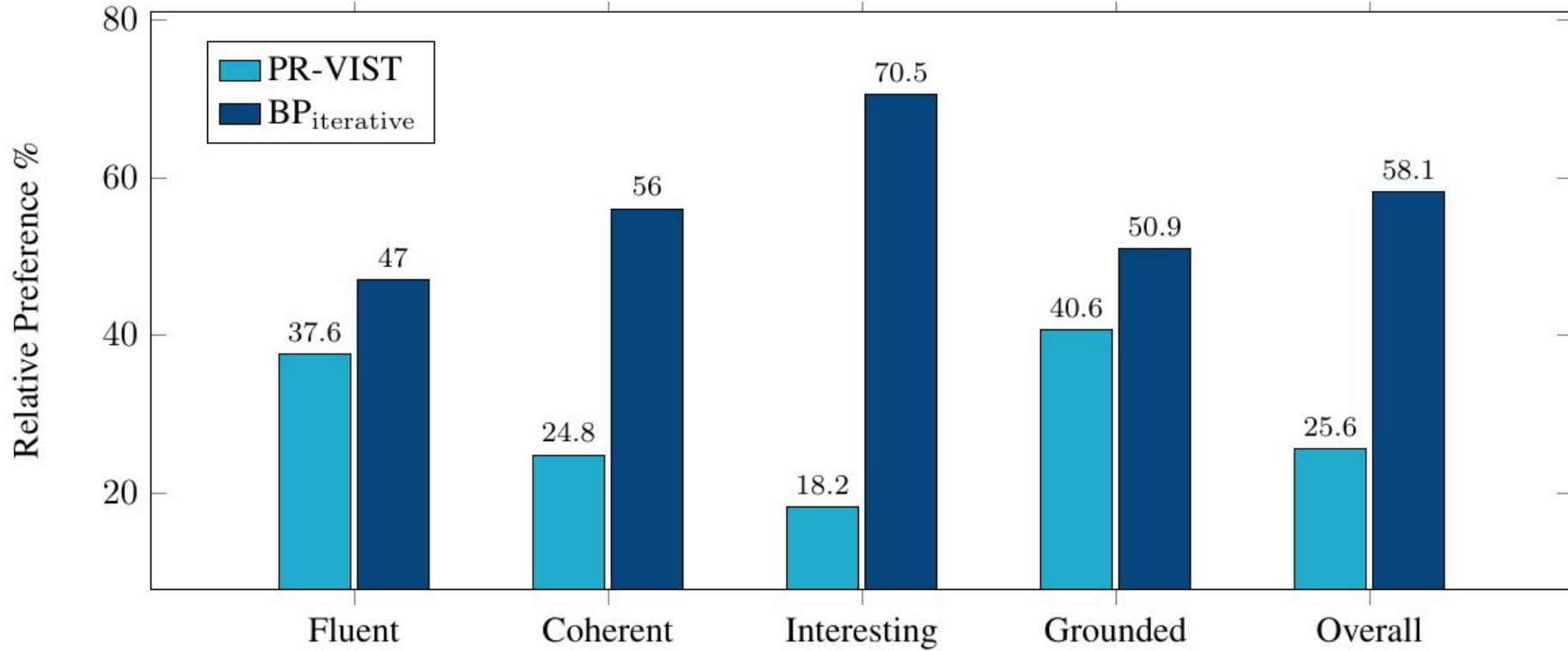
- Automatic evaluation using Blue, Rouge, Meteor is problematic – mostly measures fluency, and all modern LMs generate fluent output
- Also, a story may be interesting, coherent, and grounded in the images, but completely different from the reference
- We use **automatic evaluation** using repetition and grounding, plus Mauve for naturalness (similarity of distribution with human text)
- More importantly, we ask **human judges** to evaluate fluency, coherence, interestingness and grounding
- We compare to a range of SotA models and to GPT 3.5 (with and without Blueprint)

Automatic Evaluation



VIST is a SotA model in the literature. It doesn't use pretrained LMs; see paper for other comparisons

Human Evaluation



Example Output



PR-VIST: I bought a cart to the market with some food. It looked great in the market. There were so many of their enthusiasm. They even had a sale with their bumper. The market was sold out of fresh vegetables. the [location] car is very nice and beautiful.

VP-BART: I went to the market yesterday. There were many different kinds of fruits there. I bought a lot of them. They were very expensive. Afterward I went back home.

Iterative: Today we decided to take a small shopping trip to the Market. The market had so many wonderful things to choose from that we looked at and bought so many of them. There was an array of different fruits that I could hardly resist buying. There was also a huge array of various types of pepper. Finally the day had come to an end and we piled in the taxi back to head home.

GPT-3.5 + BP: On Saturday morning, I visited the local market. It had a lot of fruits and vegetables being sold. The atmosphere was friendly and buzzing with many vendors ready to help you. The fresh produce was the best thing about the market. You can always find friendly vendors that sell the most delicious produce. The key to finding the best bargains at the market is bargaining. The market is always a great place to shop for fresh fruits and vegetables.

Human: Shoppers arrived early for the market. There was a variety of goods sold. Some carried away large bundles. Every item was fresh and colorful. The spot was a favorite among produce shoppers.

Analysis of Results

- Pretrained language models can produce better stories than specialized models trained from scratch
- The visual prefix is an effective interface between image and text; we don't need multimodal models
- Blueprint model output is most grounded, by automatic and human evaluation
- Iterative planning strategy that generates sentence by sentence works best
- GPT-3.5 struggles with Blueprints; they reduce its performance in most metrics

Conclusions

- Blueprint-based model generates output that is more **coherent, interesting, and grounded** than existing methods
- Blueprints aid in **selecting key concepts and guiding narrative construction**
- Blueprints are **controllable**: generate longer or shorter stories (more/less iterations), emphasize entities or characters (filter Blueprint), etc.
- Blueprints are **interpretable** and could enable human-in-the-loop and personalized storytelling

[Narayan et al. Conditional Generation with a Question-Answering Blueprint. TACL, 11, 2023. 974-996](#)

[Liu et al. Visual Storytelling with Question-Answer Plans. Findings of the Association for Computational Linguistics: EMNLP 2023](#)