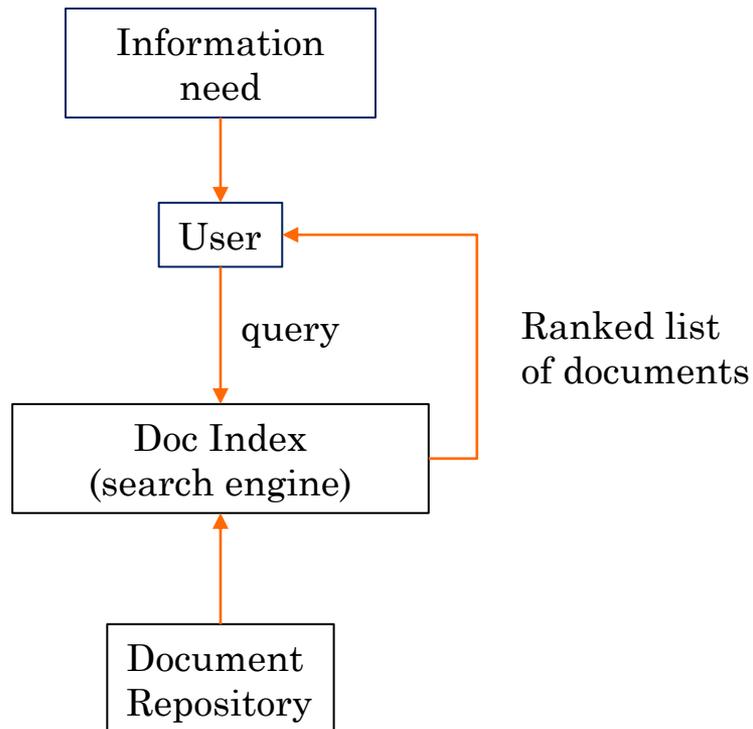


NEXT-GENERATION INFORMATION RETRIEVAL

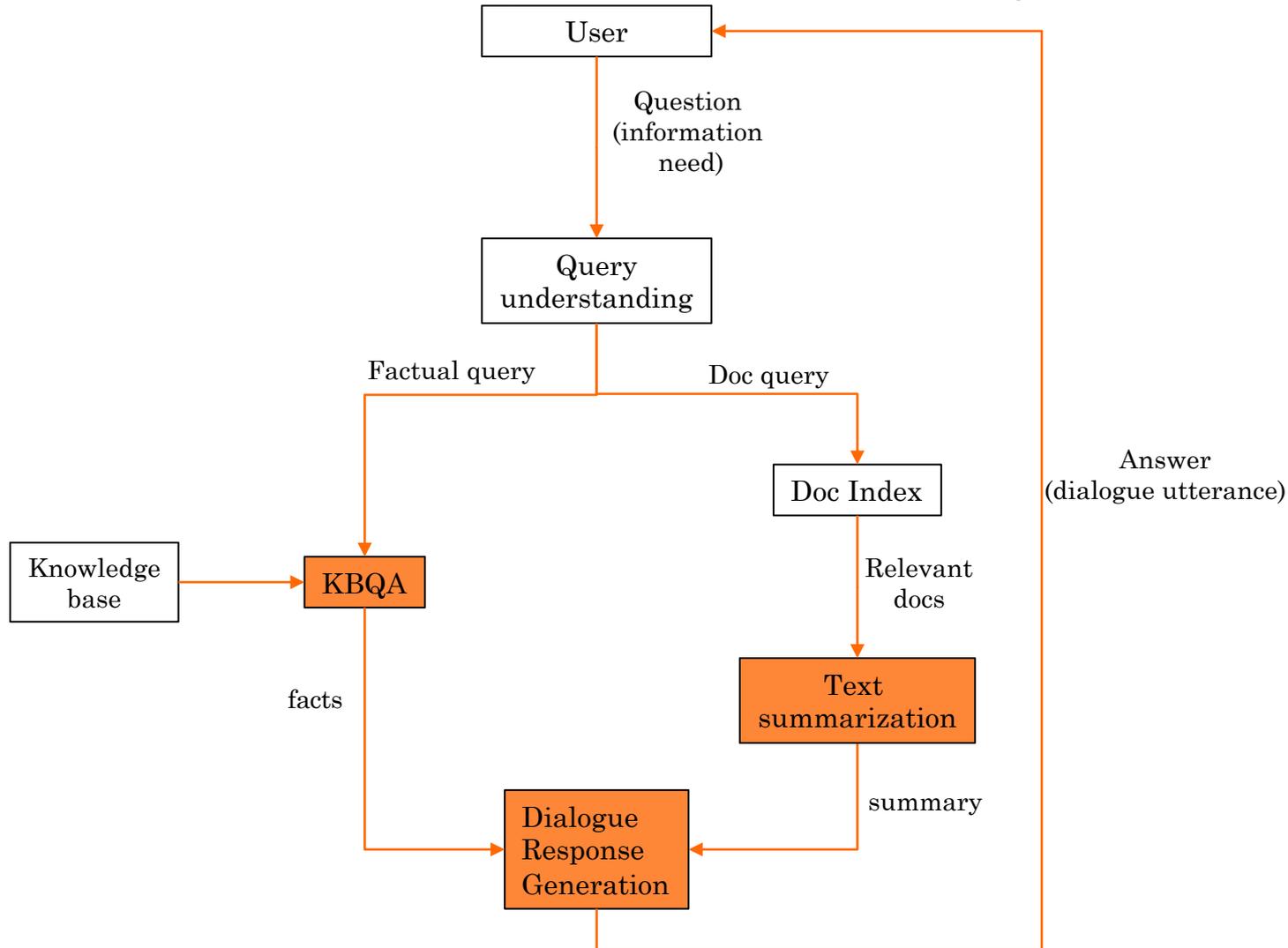
CONTENTS

- Classic vs Next-gen IR
- Preliminaries
 - Language model
 - Encoder-decoder (Seq2seq) model
- LLM: Transformer, BERT, GPT
- Text Summarization
- KBQA
- Open-domain QA (New Bing)

CLASSIC IR WORKFLOW



NEXT-GEN IR (OPEN-DOMAIN QA) WORKFLOW



KBQA (GOOGLE SEARCH)

Google X

[Images](#) [Videos](#) [News](#) [Books](#) [Maps](#) [Shopping](#) [Flights](#) [Finance](#) [Tools](#)

About 31,300,000 results (0.48 seconds)

Titanic / Director

James Cameron



James Francis Cameron CC (born August 16, 1954) is a Canadian filmmaker. A major figure in the post-New Hollywood era, he is considered one of the industry's most innovative filmmakers, regularly pushing the boundaries of cinematic capability with his use of novel technologies.

Wikipedia
https://en.wikipedia.org/wiki/James_Cameron

[James Cameron - Wikipedia](#)

People also search for

James Cameron

Canadian filmmaker

James Francis Cameron CC is a Canadian filmmaker. A major figure in the post-New Hollywood era, he is considered one of the industry's most innovative filmmakers, regularly pushing the boundaries of cinematic capability with his use of novel technologies. [Wikipedia](#)

Born: August 16, 1954 (age 68 years), Kapuskasing, Canada

Edited works: [Trade and the Environment: Bridging the Gap](#), [MORE](#)

Art directed: [Battle Beyond the Stars](#)

Productions designed: [Xenogenesis](#), [Galaxy of Terror](#)

Organizations founded: [Lightstorm Entertainment](#), [Digital Domain](#), [Cameron Pace Group](#)

Production companies: [Lightstorm Entertainment](#), [Digital Domain](#), [Cameron Pace Group](#)

Awards: [Academy Award for Best Picture](#), [MORE](#)

OPEN-DOMAIN QA (NEW BING)

Microsoft Bing

SEARCH

CHAT



I am planning a trip for our anniversary in September. What are some places we can go that are within a 3 hour flight from London Heathrow?

Congratulations on your anniversary! 🎉 There are many places you can go that are within a 3 hour flight from London Heathrow. Here are some suggestions based on your preferences and the best destinations in Europe in September ⁴ ⁵ ⁶:

- If you like **beaches** and **sunshine**, you can fly to **Malaga** in Spain ¹, which has a pleasant climate, a lively culture, and a beautiful coastline. You can enjoy the sandy beaches, the historic monuments, and the delicious tapas. Malaga is also a great base to explore other nearby attractions, such as Granada, Seville, and Cordoba.
- If you like **mountains** and **lakes**, you can fly to **Annecy** in France ⁴, which is a charming alpine town near the Swiss border. Annecy is known as the "Venice of France" because of its canals and bridges. You can admire the stunning scenery, visit the medieval castle, and try some local cheese and wine. Annecy is also close to Geneva, Chamonix, and Lyon, if you want to see more of the region.
- If you like **art** and **history**, you can fly to **Florence** in Italy ⁶, which is the birthplace of the Renaissance and a UNESCO World Heritage Site. Florence is a treasure trove of artistic and architectural masterpieces, such as the Duomo, the Uffizi Gallery, and the Ponte Vecchio. You can also explore the Tuscan countryside, taste the famous gelato, and shop for leather goods.



Ask me anything...

Feedback

QUIZ (NEW BING SEARCH)

- Besides taking the summary from retrieved documents and possible facts from KB as input, what else should the New Bing search engine take as input to generate the result that we see now?

LANGUAGE MODELS

$$P(s) = \prod_{i=1}^k P(w_i | w_1 \cdots w_{i-1})$$

○ Unigram Language Model:

- $P(w_1, w_2, \dots, w_n) = P(w_1) * P(w_2) * \dots * P(w_n)$

○ Bigram Language Model:

- $P(w_1, w_2, \dots, w_n) = P(w_1) * P(w_2 | w_1) * \dots * P(w_n | w_{n-1})$

○ Exponential Language Model:

$$P(w_m | w_1, \dots, w_{m-1}) = \frac{1}{Z(w_1, \dots, w_{m-1})} \exp(\alpha^T f(w_1, \dots, w_m))$$

- Z is a partition function (normalization), α is a params vector, f is a feature function.

LANGUAGE MODELS

- Neural network language model
 - Use continuous representation (embeddings) of words to make predictions

$$P(w_t \mid \text{context}) \forall t \in V.$$

- the network is trained to predict a probability distribution over the vocabulary, given some linguistic context.
- The context (bag of words) may be:

$$P(w_t \mid w_{t-k}, \dots, w_{t-1})$$

$$P(w_t \mid w_{t-k}, \dots, w_{t-1}, w_{t+1}, \dots, w_{t+k}).$$

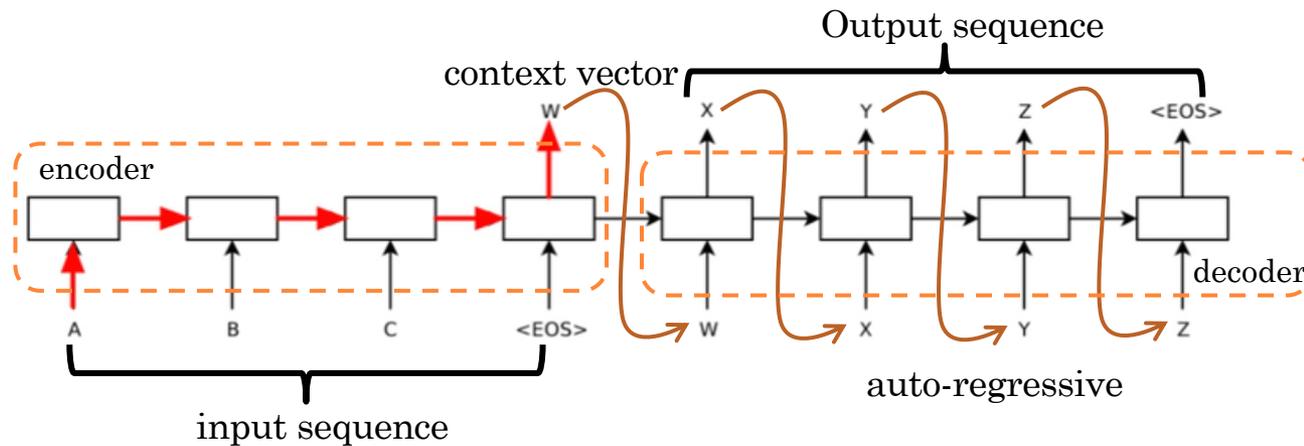
ENCODER-DECODER (SEQ2SEQ) MODEL



- Input/output are sequences of words
- State is an internal representation of the input (a vector)
- Ubiquitous in NLG applications:
 - Machine translation
 - Text summarization
 - Dialogue systems
 - Question answering
 - etc.



ENCODER-DECODER (SEQ2SEQ) MODEL

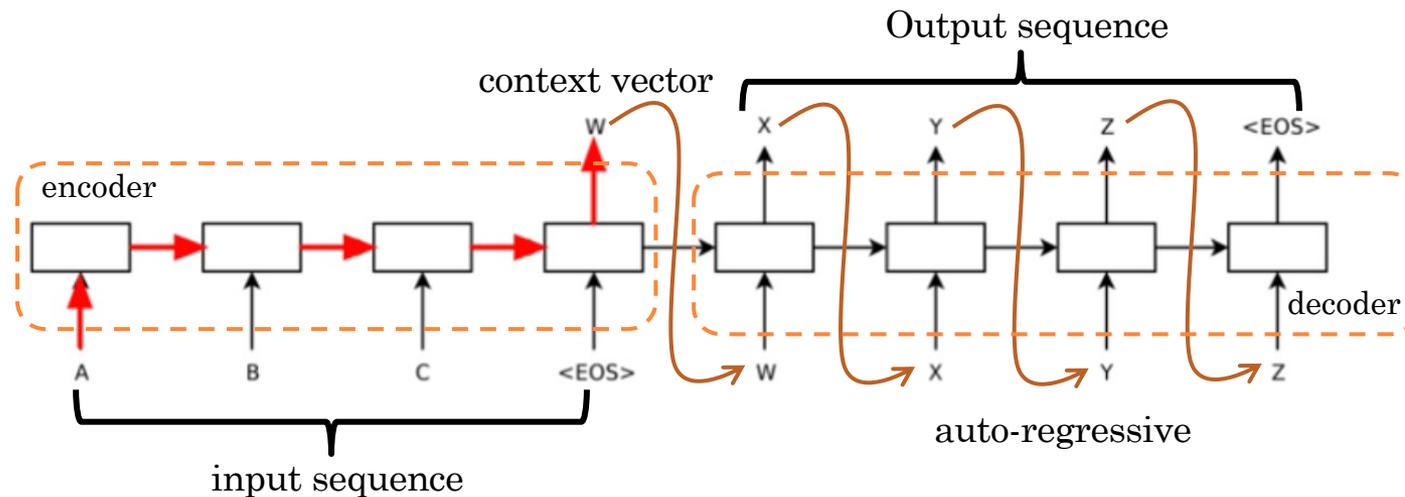


Each rectangle is a recurrent neural network (RNN) unit (hidden layer)



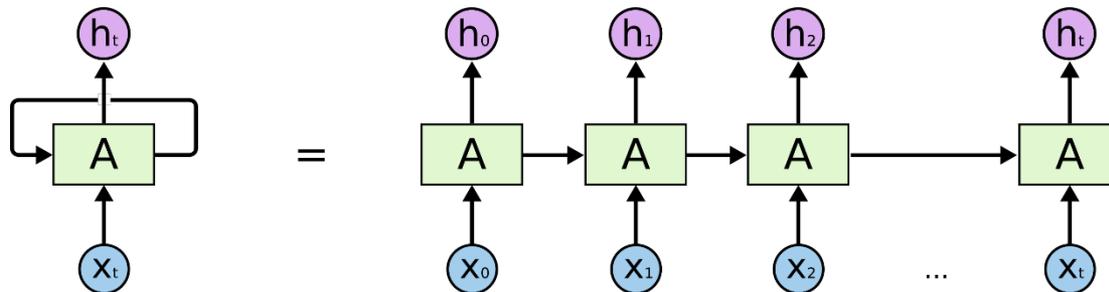
QUIZ (ENCODER-DECODER ARCHI)

- What is considered the intermediate state in this picture?

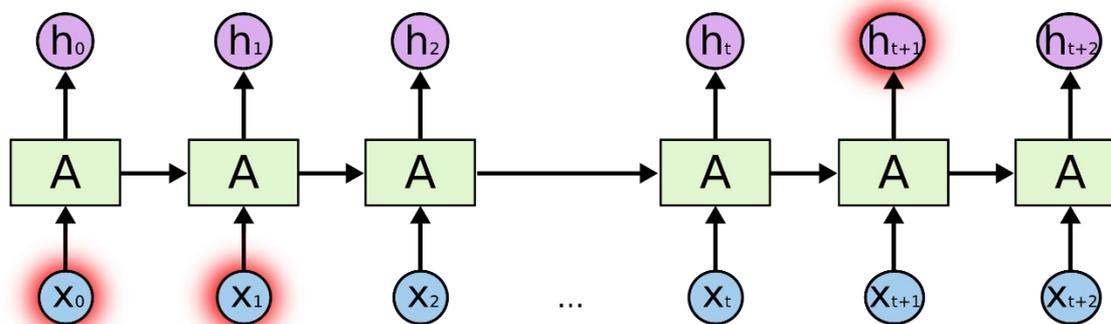


RECURRENT NEURAL NETWORK

- Basic RNN architecture:

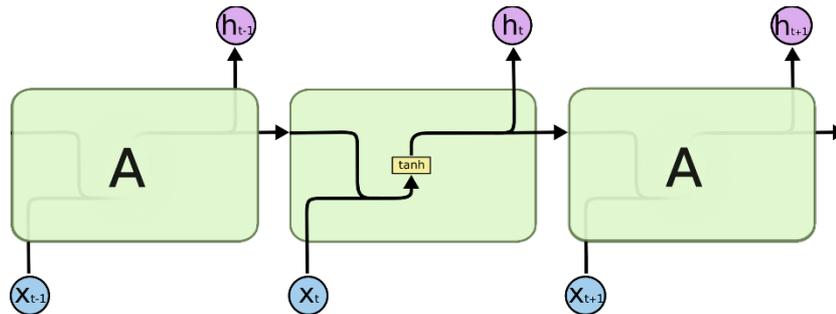


- RNN tends to forget long-term dependencies:

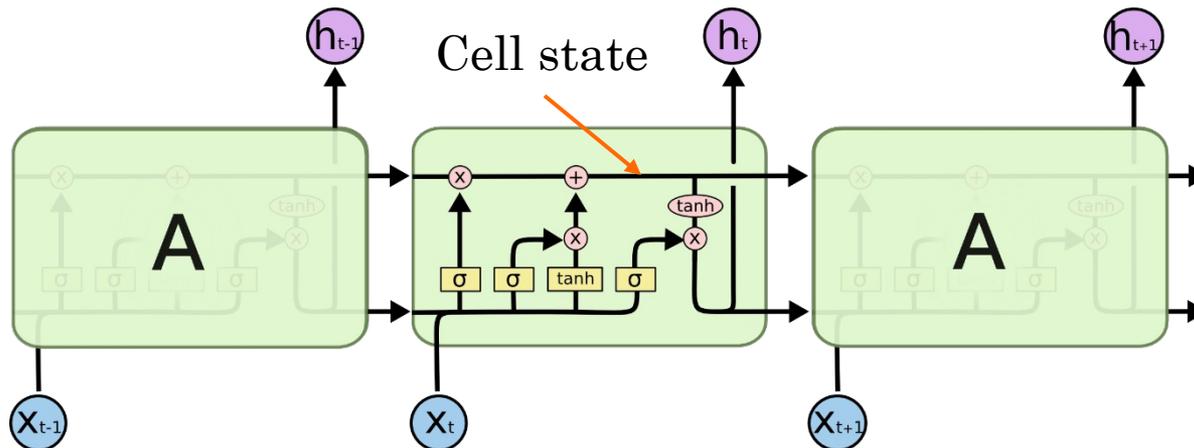


LONG SHORT-TERM MEMORY (LSTM)

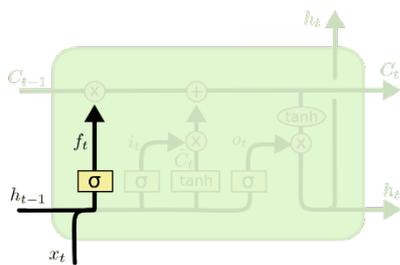
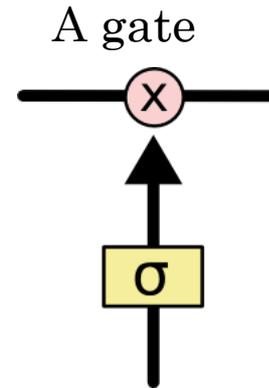
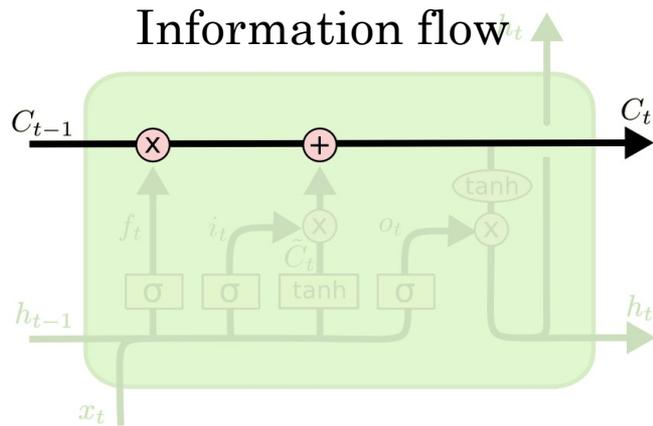
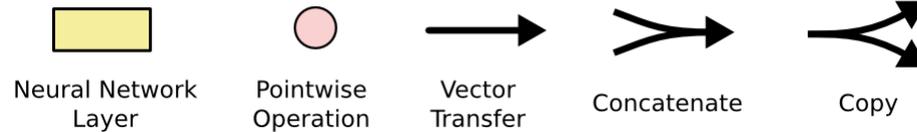
- Basic RNN unit:



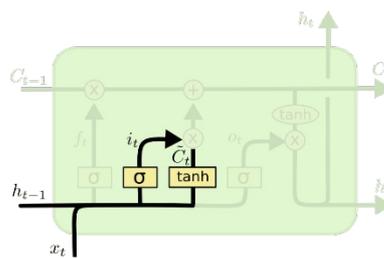
- LSTM unit (Input, output and forget gates):



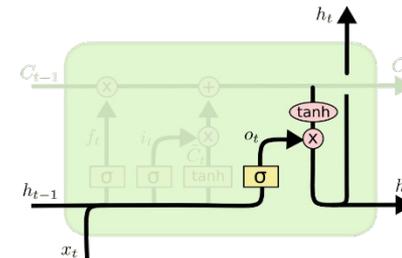
LONG SHORT-TERM MEMORY (LSTM)



forget gate



input gate



output gate

LARGE LANGUAGE MODELS

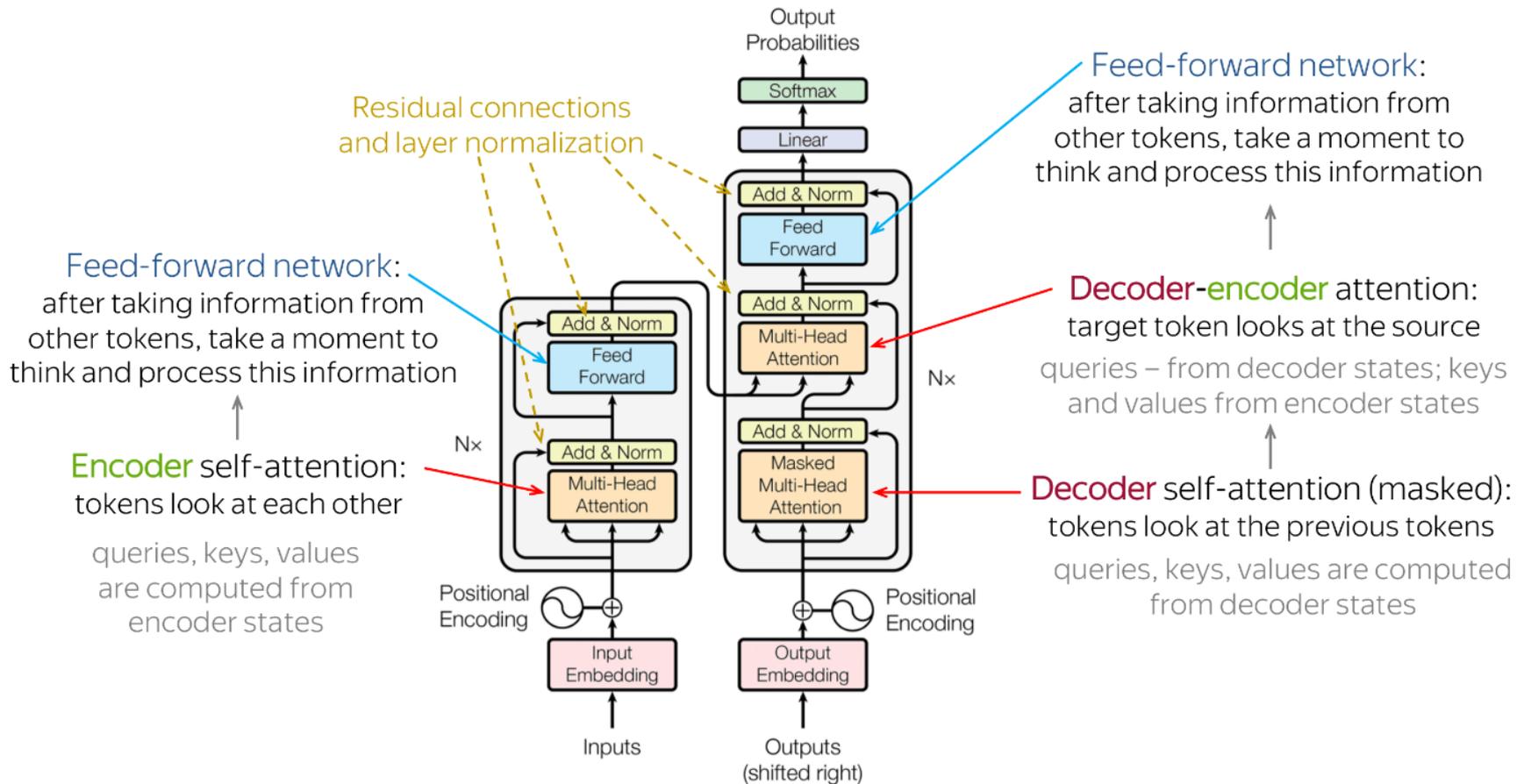
- Foundational machine learning models
- Trained by deep neural network algorithms
- Contain multiple layers of neural networks, like *recurrent* layers, *feedforward* layers, *embedding* layers, and *attention* layers
 - Embedding layer: convert a word into high-dimensional vector representation – syntactic and semantic information of words
 - Feedforward layer: non-linear transformation of the input embeddings
 - Recurrent layer: process sequence of words, long-range dependencies
 - Attention layer: focus selectively on different part of the input sequence

EXAMPLES OF LARGE LANGUAGE MODELS

- Transformer
- BERT (Bidirectional Encoder Representations from Transformers)
- GPT (Generative Pre-trained Transformer)
- XLNet
- RoBERTa (Robustly Optimized BERT Pretraining Approach)
- T5 (Text-to-Text Transfer Transformer)

TRANSFORMER

- Attention is all you need!



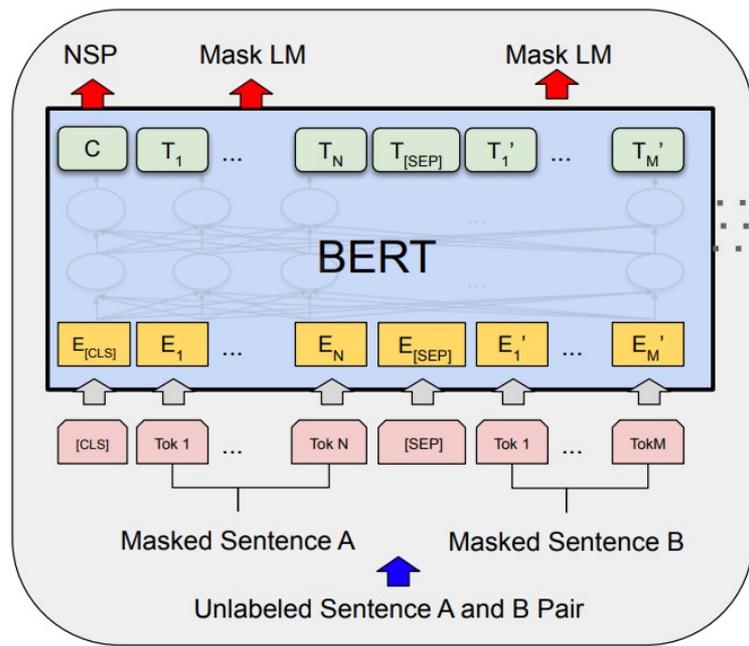
QUIZ: TRANSFORMER

- Can you think of a disadvantage of the transformer encoder?

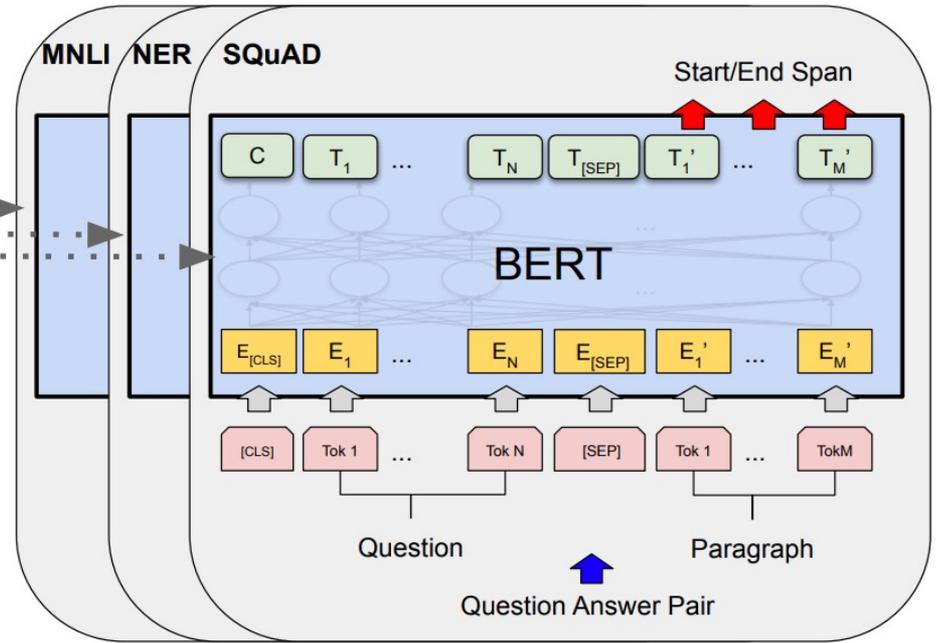
BERT

- Transformer encoder only
- For classification, not generation
- Pretrained on unlabeled text corpus
 - Task 1: Masked language model (MLM)
 - Task 2: Next sentence prediction (NSP)
- Fine-tuned on task-specific data

BERT



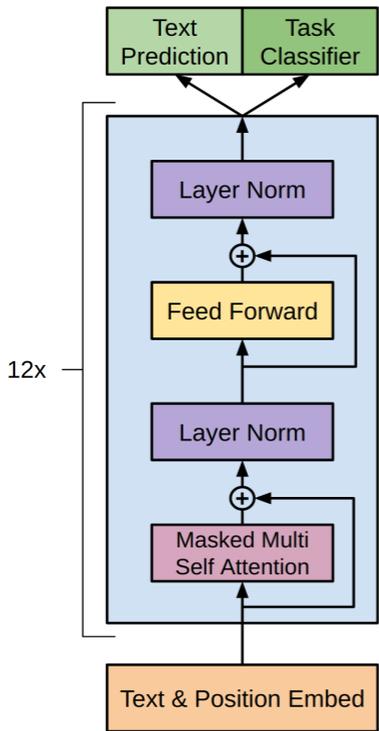
Pre-training



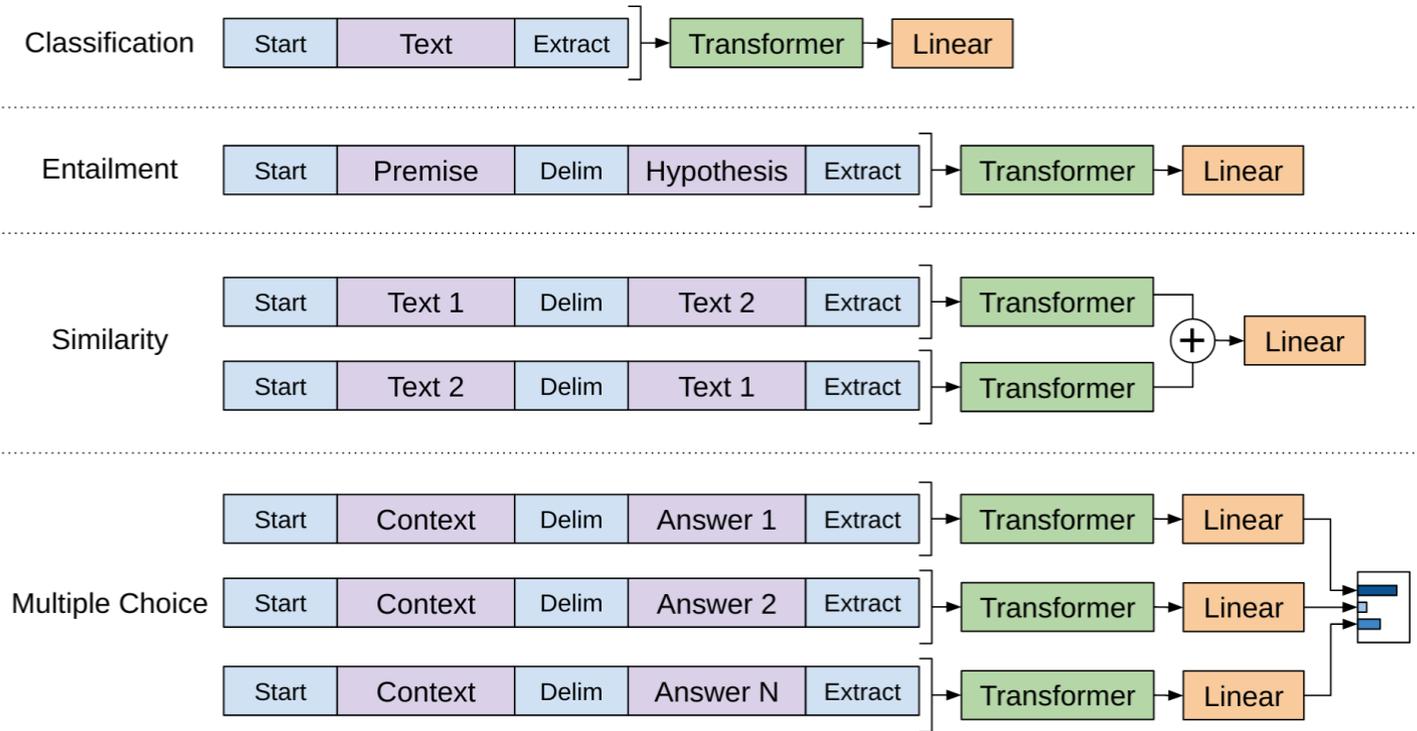
Fine-Tuning

50% of the tokens masked for MLM
50% of the next sentence is true, other 50%
is random

GPT



Pretrain



Fine-tune

GPT

Model	Architecture	Parameter count	Training data	Release date
GPT-1	12-level, 12-headed Transformer decoder (no encoder)	117 million	<u>BookCorpus</u> : 4.5 GB of text, from 7000 unpublished books.	June 11, 2018
<u>GPT-2</u>	GPT-1, but with modified normalization	1.5 billion	WebText: 40 GB of text, 8 million documents, from 45 million webpages upvoted on Reddit.	February 14, 2019
<u>GPT-3</u>	GPT-2, but with modification to allow larger scaling	175 billion	570 GB plaintext, 0.4 trillion tokens. Mostly CommonCrawl, WebText, English Wikipedia, and two books corpora	June 11, 2020
<u>GPT-4</u>	Also trained with both text prediction and RLHF; accepts both text and images as input.	Undisclosed	Undisclosed	March 14, 2023

TEXT SUMMARIZATION

- Text summarization is classified into two types — **Extractive** and **Abstractive Summarization**.
- **Extractive Summarization:** The extractive text summarization process extracts the main points of a text without any alteration to those points and rearranging the order of that points and the grammar to get the soul out of the summary.
- **Abstractive Summarization:** The Abstractive methods use advanced techniques to get a whole new summary. Some parts of this summary might not even appear within the original text.

SUMMARIZATION DATASET

○ CNN/DailyMail (312k instances)

```
{'id': '0054d6d30dbcad772e20b22771153a2a9cbeaf62',  
  'article': '(CNN) -- An American woman died aboard a cruise ship that docked at Rio de Janeiro on Tuesday, the same ship on which 86 passengers previously fell ill, according to the state-run Brazilian news agency, Agencia Brasil. The American tourist died aboard the MS Veendam, owned by cruise operator Holland America. Federal Police told Agencia Brasil that forensic doctors were investigating her death. The ship's doctors told police that the woman was elderly and suffered from diabetes and hypertension, according the agency. The other passengers came down with diarrhea prior to her death during an earlier part of the trip, the ship's doctors said. The Veendam left New York 36 days ago for a South America tour.'  
  'highlights': 'The elderly woman suffered from diabetes and hypertension, ship's doctors say .\nPreviously, 86 passengers had fallen ill on the ship, Agencia Brasil says .'
```

○ Xsum (226k instances)

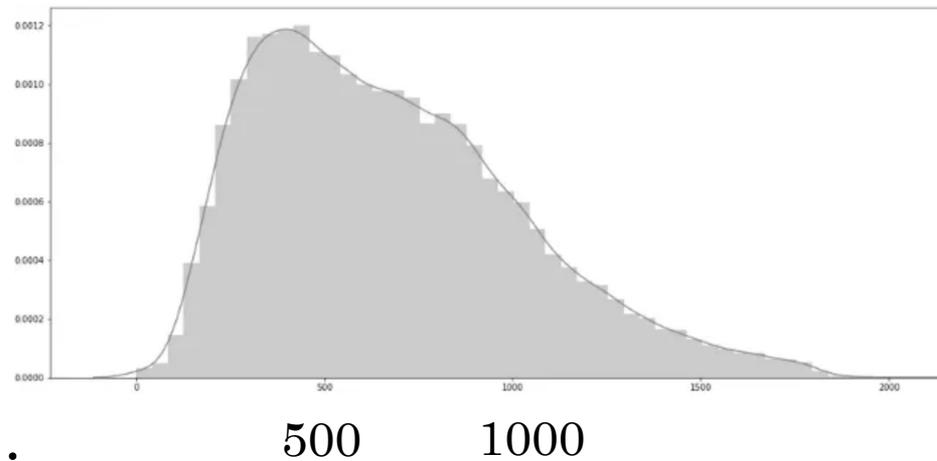
"Belgian cyclist Demoitie died after a collision with a motorbike during Belgium's Gent-Wevelgem race. The 25-year-old was hit by the motorbike after several riders came down in a crash as the race passed through northern France. "The main issues come when cars or motorbikes have to pass the peloton and pass riders," Team Sky's Rowe said. "That is the fundamental issue we're looking into. "There's a lot of motorbikes in and around the race whether it be cameras for TV, photographers or police motorbikes. "In total there's around 50 motorbikes that work on each race. "We've got a riders union and we're coming together to think of a few ideas, whether we cap a speed limit on how fast they can

"Welsh cyclist Luke Rowe says changes to the sport must be made following the death of Antoine Demoitie."

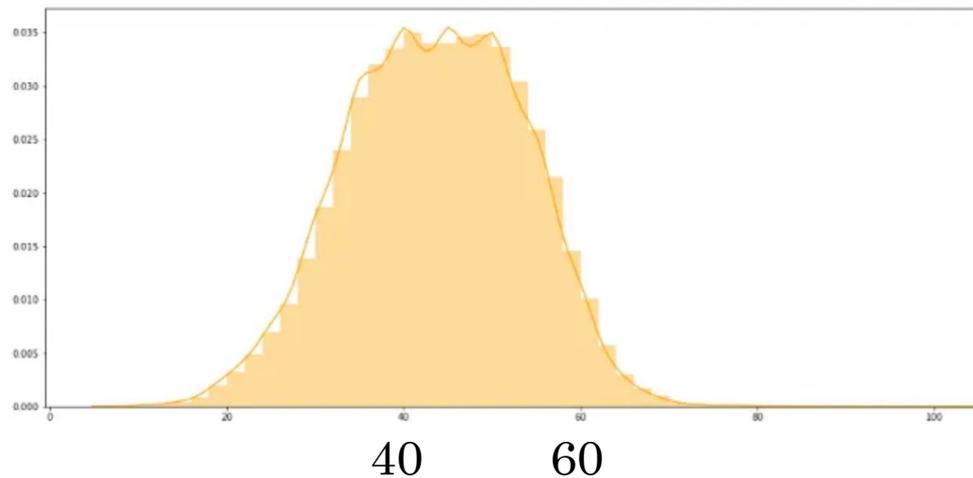
"35932467"

CNN/DM DATASET

○ Article length:



○ Summary length:



EVALUATION METRIC

○ BLEU Score

- “bilingual evaluation understudy”
- BLEU scores range from 0 and 1.
- If predicted and original text is a similar score close to 1 and vice-versa.

○ ROUGE Score

- “Recall-Oriented Understudy for Gisting Evaluation”
- ROUGE-1 refers to the overlap of unigram (each word) between the system and reference summaries.
- ROUGE-2 refers to the overlap of bigrams between the system and reference summaries.
- ROUGE-L: Longest Common Subsequence (LCS) based statistics. The longest common subsequence problem takes into account sentence-level structure similarity naturally and identifies the longest cooccurring in sequence n-grams automatically.

EVALUATION METRICS

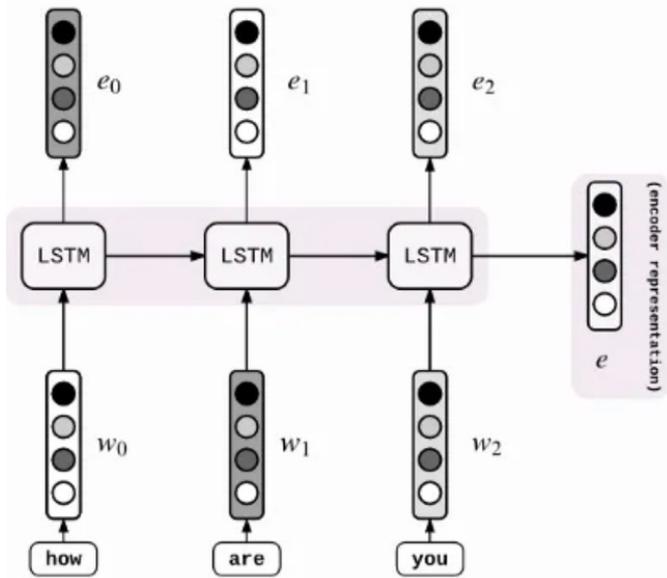
$$\text{ROUGE recall} = \frac{\text{number of overlapping words}}{\text{total words in reference summary}},$$

$$\text{ROUGE precision} = \frac{\text{number of overlapping words}}{\text{total words in system summary}},$$

- Normally we present ROUGE-F1 scores, which is calculated as we learned before.

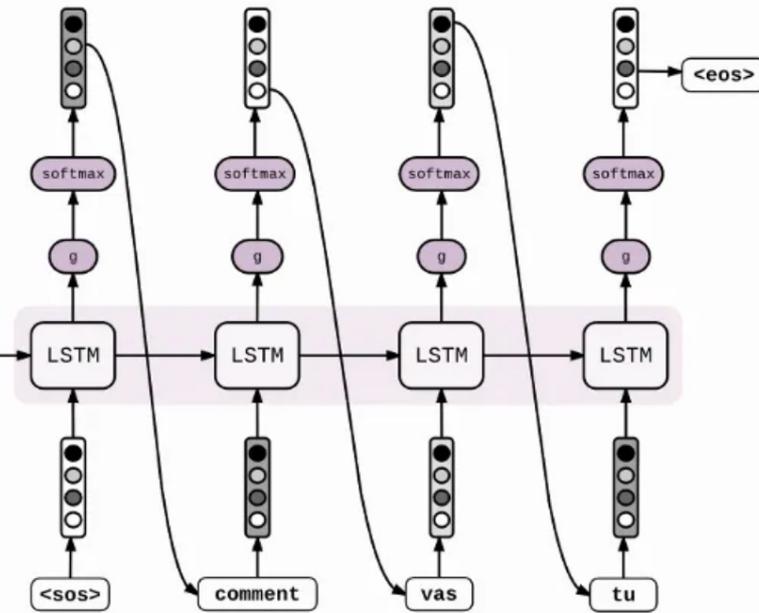
A BASELINE LSTM APPROACH

@Training time: loss is computed between predicted word from prev ref word and the target ref word.



Input Doc

Predicted Summary



Reference Summary

BASELINE PREDICTED SUMMARIES

Original summary: `_START_ workers waste little time removing images of ro
d blagojevich from state capitol illinois senate voted 59 0 on thursday to
remove blagojevich from office blagojevich was arrested on corruption char
ges continues to proclaim innocence _END_`

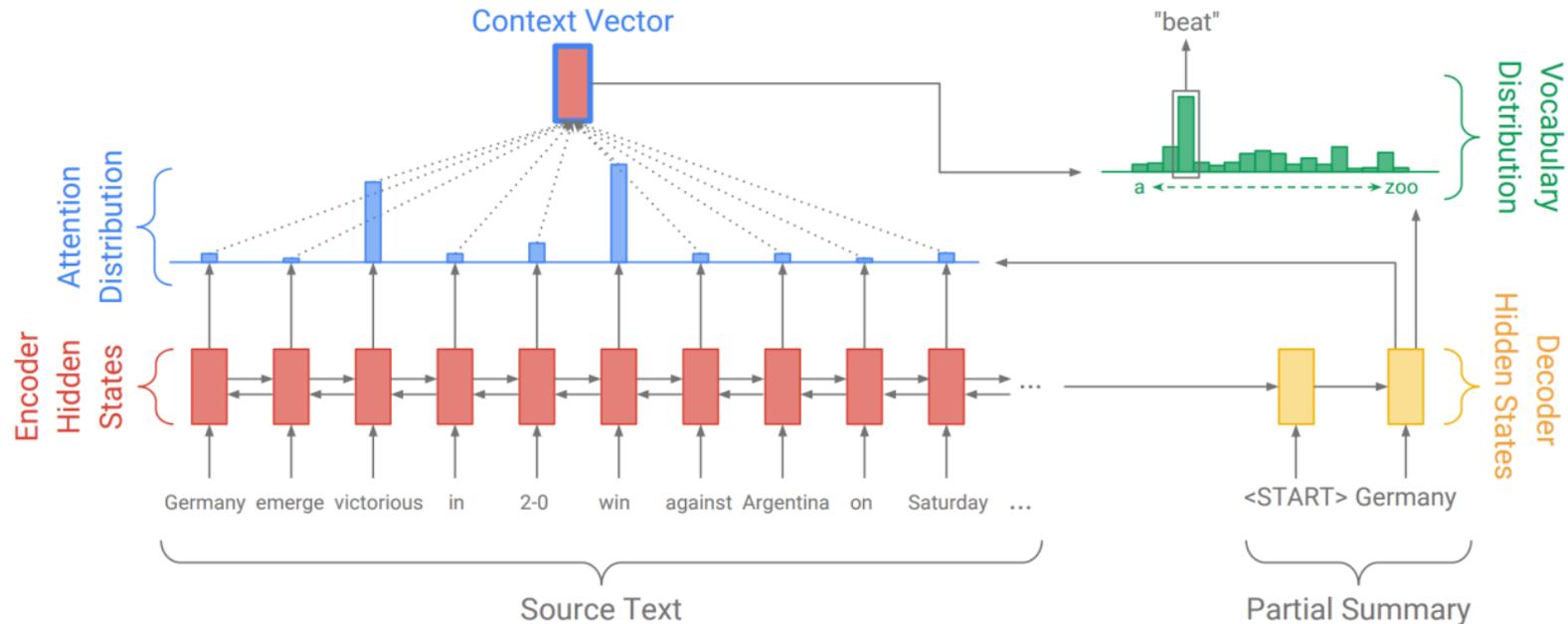
Predicted summary: `to be nameplate to nameplate nameplate nameplate namep
late nameplate nameplate nameplate nameplate nameplate nameplate nameplate
nameplate nameplate nameplate nameplate nameplate nameplate nameplate name
plate nameplate nameplate nameplate nameplate and nameplate nameplate 59`

Original summary: `_START_ andrzej lepper was a former deputy prime minist
er and agriculture minister a populist he founded his trade union and poli
tical party to help suffering farmers his self defense party paid tribute
to a distinguished politician statesman _END_`

Predicted summary: `to lepper lepper lepper lepper lepper lepper lepper le
pper lepper
lepper lepper and andrzej lepper andrzej lepper andrzej lepper andrzej lep
per andrzej lepper andrzej lepper andrzej lepper andrzej lepper andrzej le
pper`

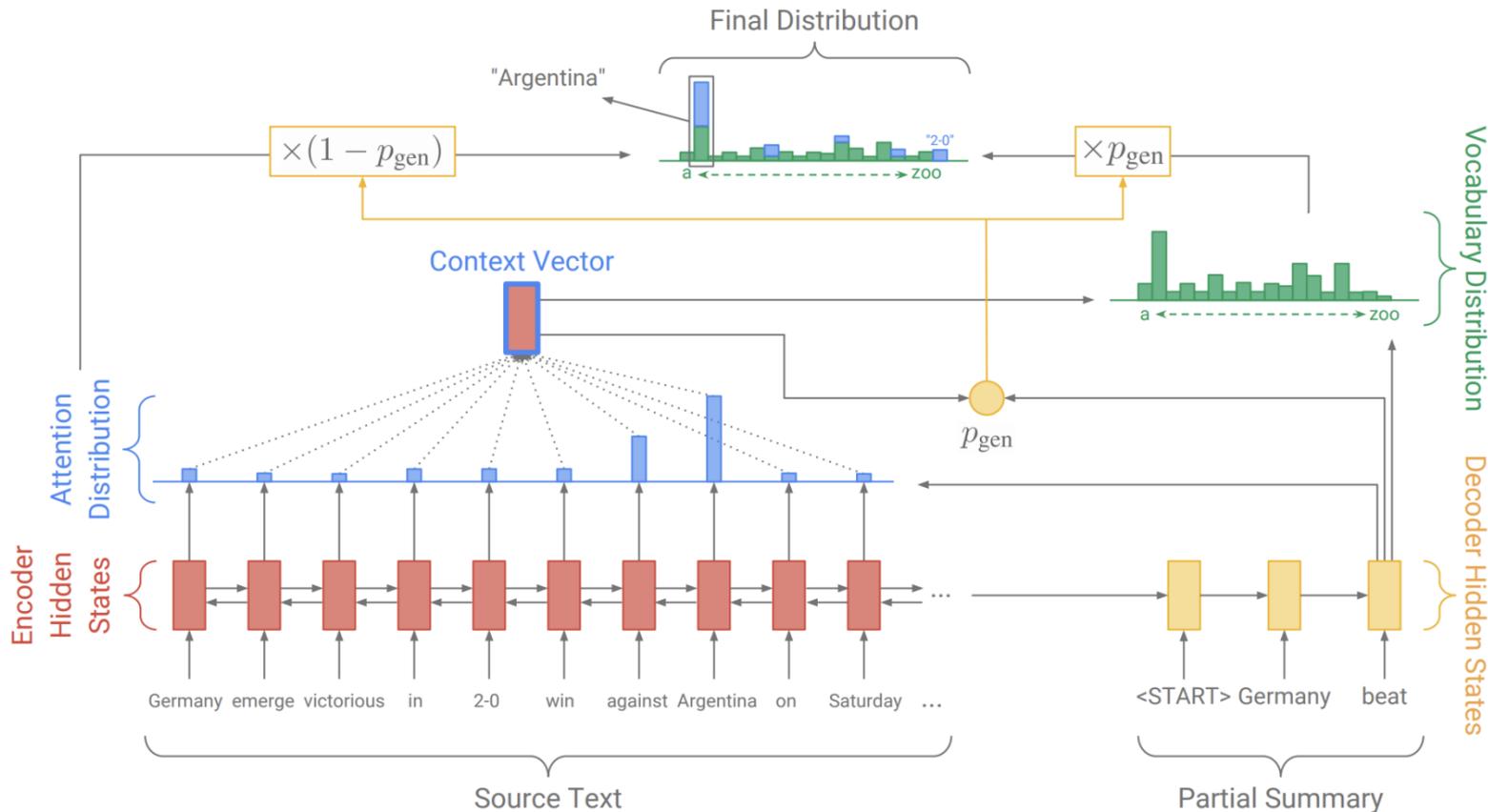
Summary is non-sensical and repeating the same words

ADDING ATTENTION



Cross attention mechanism allows the model to focus and place more “Attention” on the relevant parts of the input sequence as needed.

ADDING POINTER GENERATOR



$$P_{\text{final}}(w) = p_{\text{gen}} P_{\text{vocab}}(w) + (1 - p_{\text{gen}}) \sum_{i:w_i=w} a_i .$$

ADDING COVERAGE

- To alleviate repetition problem, we add a coverage vector:

$$c^t = \sum_{t'=0}^{t-1} a^{t'}$$

- c^t is the coverage vector at decoding timestep t , and a^t is the attention distribution at timestep t .
- Intuitively, c^t is a distribution over the source document words that represents the degree of coverage that those words have received from the attention mechanism so far.

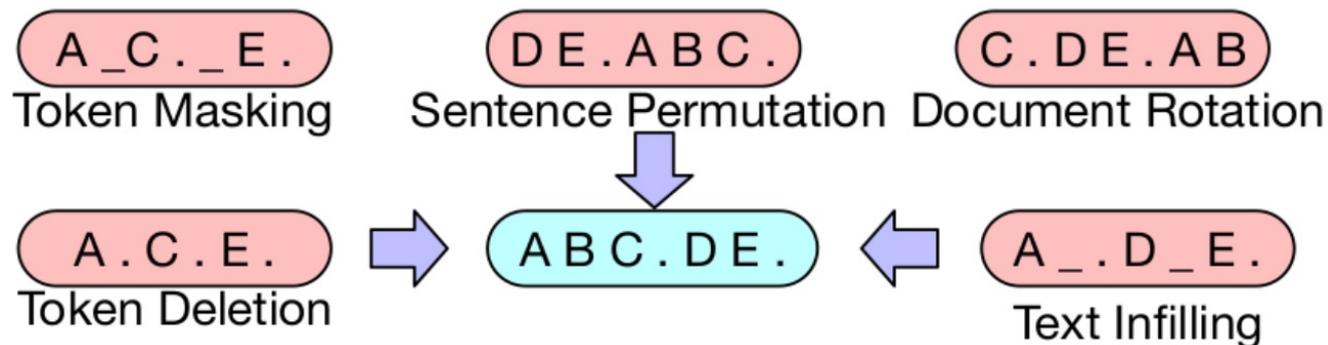
RESULTS

- Model[1] is LSTM with attention
- Model[2] is LSTM with attention and pointer-generator
- Model[3] is LSTM with attention, pointer generator and coverage

Reference	once a super typhoon , maysak is now a tropical storm with 70 mph winds . it could still cause flooding , landslides and other problems in the philippines .
Model [1]	[UNK] gained super typhoon status thanks to its sustained 150 mph winds. it 's now classified as a tropical storm. it 's expected to make landfall sunday on the southeastern coast of [UNK] province .
Model [2]	tropical storm maysak approached the asian island nation saturday . it's now classified as a tropical storm , according to the philippine national weather service . it's now classified as a tropical storm , according to the philippine weather service .
Model [3]	just a few days ago , maysak gained super typhoon status thanks to its sustained 150 mph winds . it 's now classified as a tropical storm , according to the philippine national weather service .

BART

- “Bidirectional auto-regressive transformer”
- Uses standard transformer encoder-decoder architecture
- 3 key ideas:
 - bidirectional encoding (like BERT)
 - auto-regressive decoding (like GPT)
 - noising transformation for pretraining:



CASE STUDY (A STUDENT ESSAY)

Dear @CAPS1, @CAPS2 is in my firm belief that to many people spend far too much time on their computers these days. Computers can be a very useful tool, but they can also be very harmful. I think that being on a computer consistently (day after day), degrades a person's communication skills, @CAPS2 lowers the exercise they get which hurts their healthy, but if a person manages to balance this time equally betwenn being on a computer and everything else in life then you will succed. Although computers are very valuable resources, spending too much time on them could hurt a person's communication skills. For example, if you sit on your butt all day in front of a computer you won't have any interaction with the outside world. So the next time you actually talk (and not just e-mail or text) someone, you will flounds. You won't know what to say or how to act because you haven't done @CAPS2 for so long. Now I understand that in some jobs you have to work on a computer all day. That's fine @CAPS2's just when you do @CAPS2 nepeditivley that @CAPS2 starts to harm you. You won't be able to laugh and joke around with your friends the way you used to because you'll be spending the whole time thinking about what to say next. Typing to your friends on your computer is much different than actually greating to them in person. These reasons are why I seriously believe that spending time on a computer will will your communication skills. No matter the you are on the computer. Besides hurting your communication skills. Spending too much time on the computer gives you no. Having no exercise is very unhealthy and can lead people to becoming obese. Another harmful is sitting around all day, only exercising your fingers can lead to arthritis which is becoming a bigger problems in young people not only can you get arthritis, but there are many more dangerous symtoms of having on your computer for too long. The main problem is in the lack of of health that you develop and we all know how important our health is. @CAPS2 can affect our progress how we and just lives in general. So spending to much time on a computer is a serious health issue because of the lack of exercise that you receive. Though not all is bad with the computer. @CAPS2 seem very valuable tool and is a great source of knowledge. Only be our time equally between the time we spend on the computer and the time we give to everything else in the world. Do we achieve. For instance maybe you want to look up the symtons of mans. Okay spend about an hour on the computer in searching and then go out side and play with your dog in the bright @DATE1 day. You could even combine your time on the computer with a friend. Play a video game together joke and laugh. So computers are in great thing, but we must be able to balance our time on them and on everything around us. In conclusion I say again that a computer is a great tool, only when @CAPS2 is not abused. People cannot spend too much time on the computer because @CAPS2 degrades their communication skills and ruins their health because @CAPS2 allows no exercise, however when we balance our time between the computer and everything else we got the best of both world message.

BART vs. T5 vs. PEGASUS

BART (abstractive)

Computers can be a very useful tool, but they can also be very harmful. Being on a computer consistently (day after day) degrades a person's communication skills. Having no exercise is very unhealthy and can lead people to becoming obese.

T5 (abstractive)

computers can be a very useful tool, but they can also be very harmful . being on a computer consistently (day after day), degrades a person's communication skills . people cannot spend too much time on the computer because @CAPS2 allows no exercise .

PEGASUS (abstractive)

I think that being on a computer consistently (day after day), degrades a person's communication skills, @CAPS2 lowers the exercise they get which hurts their healthy, but if a person manages to balance this time equally between being on a computer and everything else in life then you will abused. Although computers are very

PEGASUS finetuned on CNNDM news (abstractive)

People cannot spend too much time on the computer because @CAPS2 degrades their communication skills and ruins their health . Having no exercise is very unhealthy and can lead people to becoming obese . Only be our time equally between the time we spend on the computer and the time we give to everything else in the world .'

GPT-3 VARIANTS

davinci

In summary, spending too much time on a computer can be detrimental to a person's communication skills and health. It can lead to a lack of exercise, which can cause obesity and arthritis. However, computers can be a valuable tool when used in moderation and balanced with other activities. People should strive to find a balance between time spent on the computer and time spent on other activities in order to get the best of both worlds.

curie

This essay is about how computers can be a very useful tool, but can also be harmful if people spend too much time on them. The author believes that people should spend equal time on computers and everything else in their lives in order to have the best of both worlds.

babbage

1. Computers can be a great tool, but they can also be harmful.
2. Spending too much time on a computer can hurt a person's communication skills.
3. A lack of exercise can lead people to become obese.
4. A computer is a great tool, but we must be able to balance our time on it and on everything around us.

ada

The author believes that spending time on a computer for too long can harm a person's communication skills. Spending too much time on a computer can also lead to obesity. **Another harmful** is sitting around all day, only exercising their fingers can lead to arthritis which is becoming a bigger problem in young people. The author believes that spending time on a computer for too long is a health issue and that it cannot be balanced with other activities.

LIMITATIONS

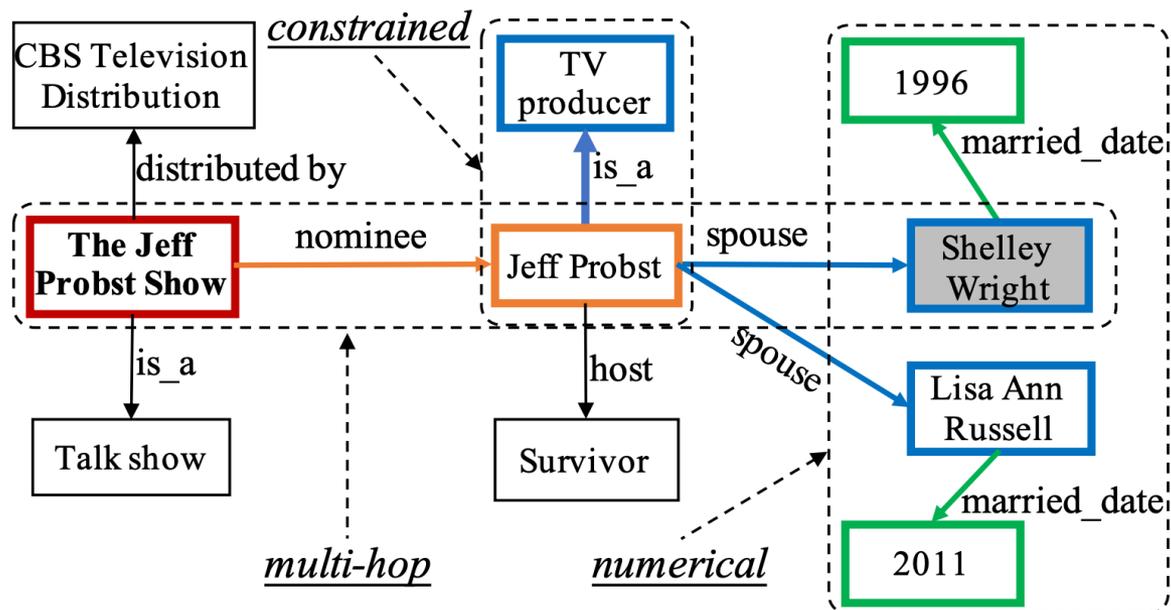
- Can't handle very long input texts
- Can't control the summary length very well
- Sometimes summary contains non-facts or hallucinations

QUIZ: MULTI-DOC SUMMARIZATION

- We have covered single document summarization, what shall we do if we want to summarize multiple documents?

COMPLEX KBQA

- A knowledge base is a graph containing edges (*subject, relation, object*)
- A question such as:
 - “Who is the first wife of TV producer that was nominated for The Jeff Probst Show?”
 - Answer: Shelley Wright



CHALLENGES OF COMPLEX KBQA

- Multi-hops
- Constrained relations
- Numerical operations
- Combinations of the above

BENCHMARK DATASETS

Datasets	KB	Size	LF	NL
WebQuestions [Berant <i>et al.</i> , 2013]	Freebase	5,810	No	No
ComplexQuestions [Bao <i>et al.</i> , 2016]	Freebase	2,100	No	No
WebQuestionsSP [Yih <i>et al.</i> , 2016]	Freebase	4,737	Yes	Yes
ComplexWebQuestions [Talmor and Berant, 2018]	Freebase	34,689	Yes	Yes
QALD series [Lopez <i>et al.</i> , 2013]	DBpedia	-	Yes	Yes
LC-QuAD [Trivedi <i>et al.</i> , 2017]	DBpedia	5,000	Yes	Yes
LC-QuAD 2.0 [Dubey <i>et al.</i> , 2019]	DBpedia, Wikidata	30,000	Yes	Yes
MetaQA Vanilla [Zhang <i>et al.</i> , 2018]	WikiMovies	400k	No	No
CFQ [Keysers <i>et al.</i> , 2020]	Freebase	239,357	Yes	No
GrailQA [Gu <i>et al.</i> , 2020]	Freebase	64,331	Yes	Yes
KQA Pro [Shi <i>et al.</i> , 2020]	Wikidata	117,970	Yes	Yes

LF: Logical Forms

NL: Rewrite LF in Natural Language

EVALUATION METRICS

○ Reliability

- Precision, Recall and F1
- Hits@1 $\text{Hits@1} = \mathbb{I}(\tilde{a}_q \in \mathcal{A}_q),$

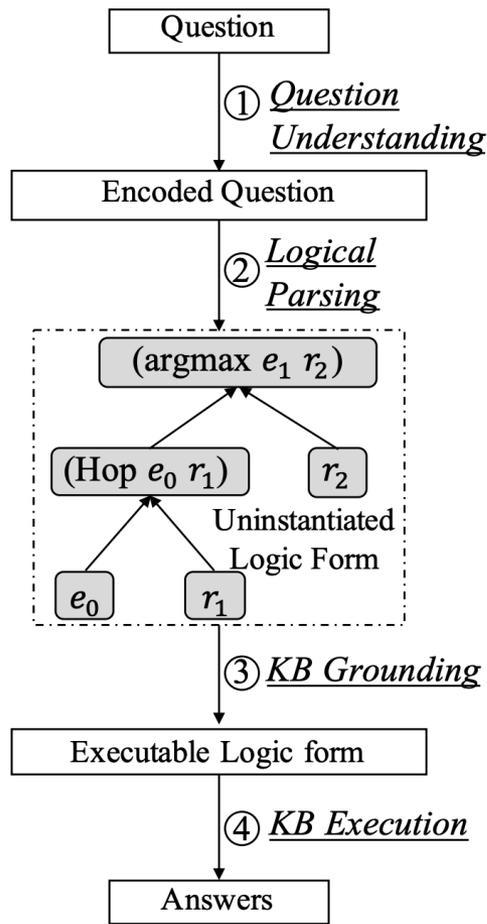
where \tilde{a}_q is the top 1 prediction in $\tilde{\mathcal{A}}_q$.

○ Robustness

- GrailQA dataset (Gu et al.)
- three levels of generalization: *i.i.d.*, *compositional*, *zero-shot*

○ System-user interaction

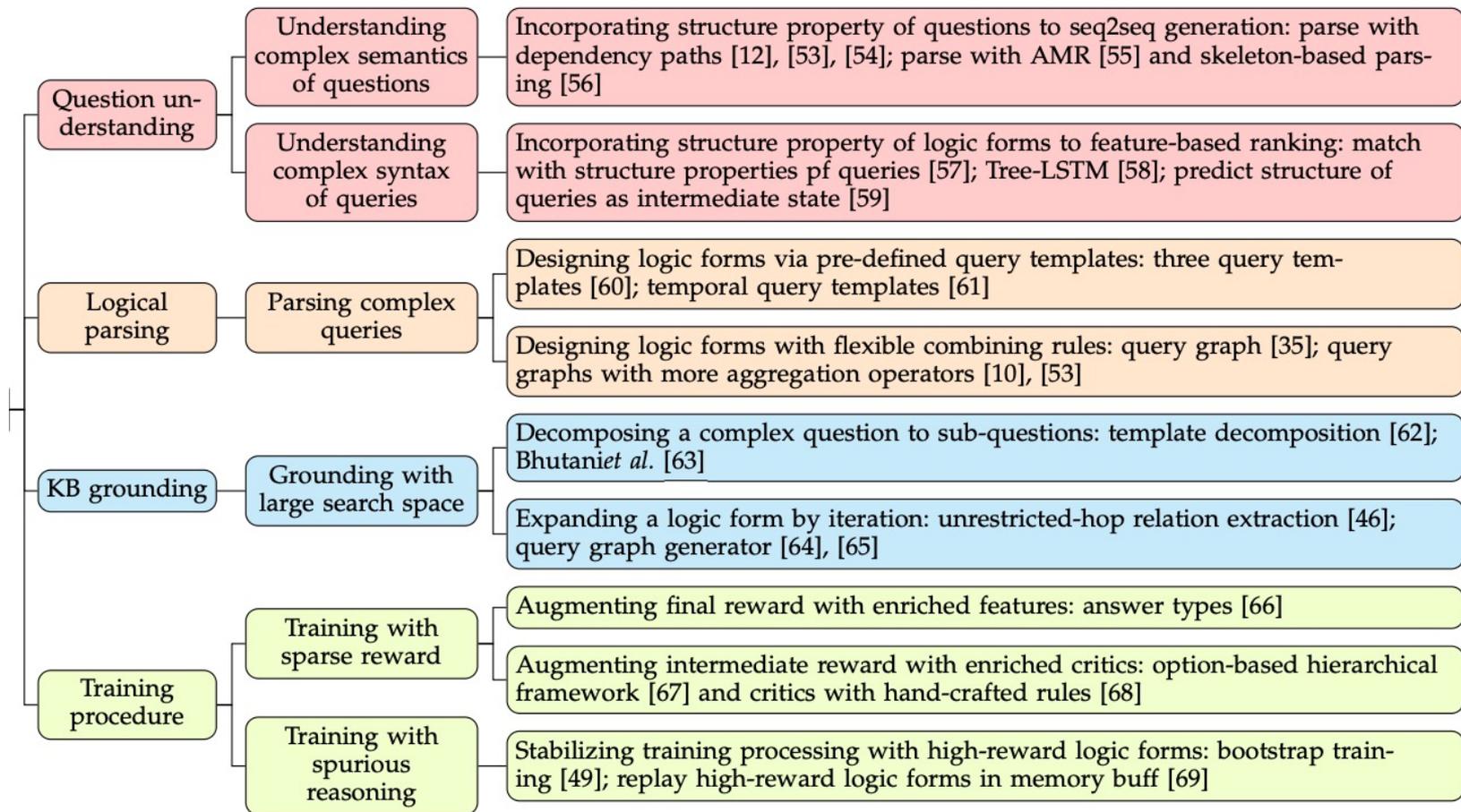
SEMANTIC PARSING APPROACH



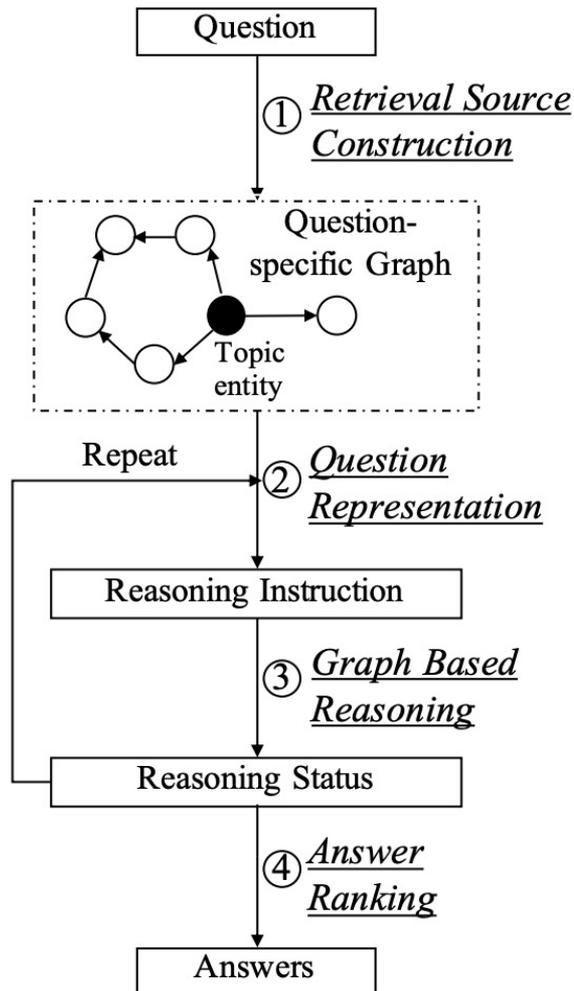
SP based-methods

- This category of methods aims at parsing a natural language utterance into logic forms. They predict answers via the following steps:
 1. Parse the natural language question into an uninstantiated logic form (e.g. a **SPARQL** query template), which is a syntactic representation of the question without the grounding of entities and relations.
 2. The logic form is then instantiated and validated by conducting some semantic alignments to structured KBs via KB grounding (obtaining, for example, an executable SPARQL query).
 3. The parsed logic form is executed against KBs to generate predicted answers.

CHALLENGES & SOLUTIONS



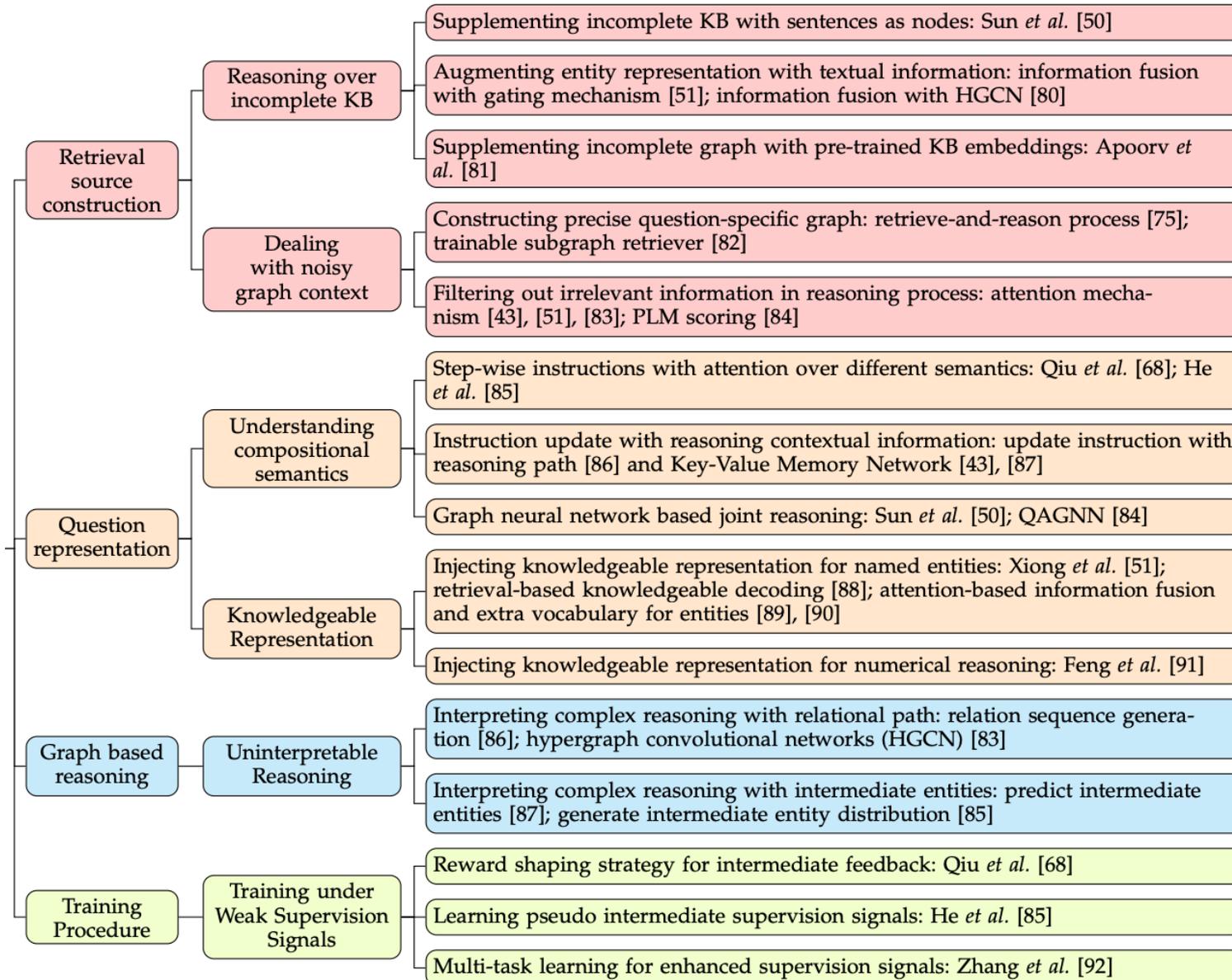
INFORMATION RETRIEVAL BASED APPROACH



IR based-methods

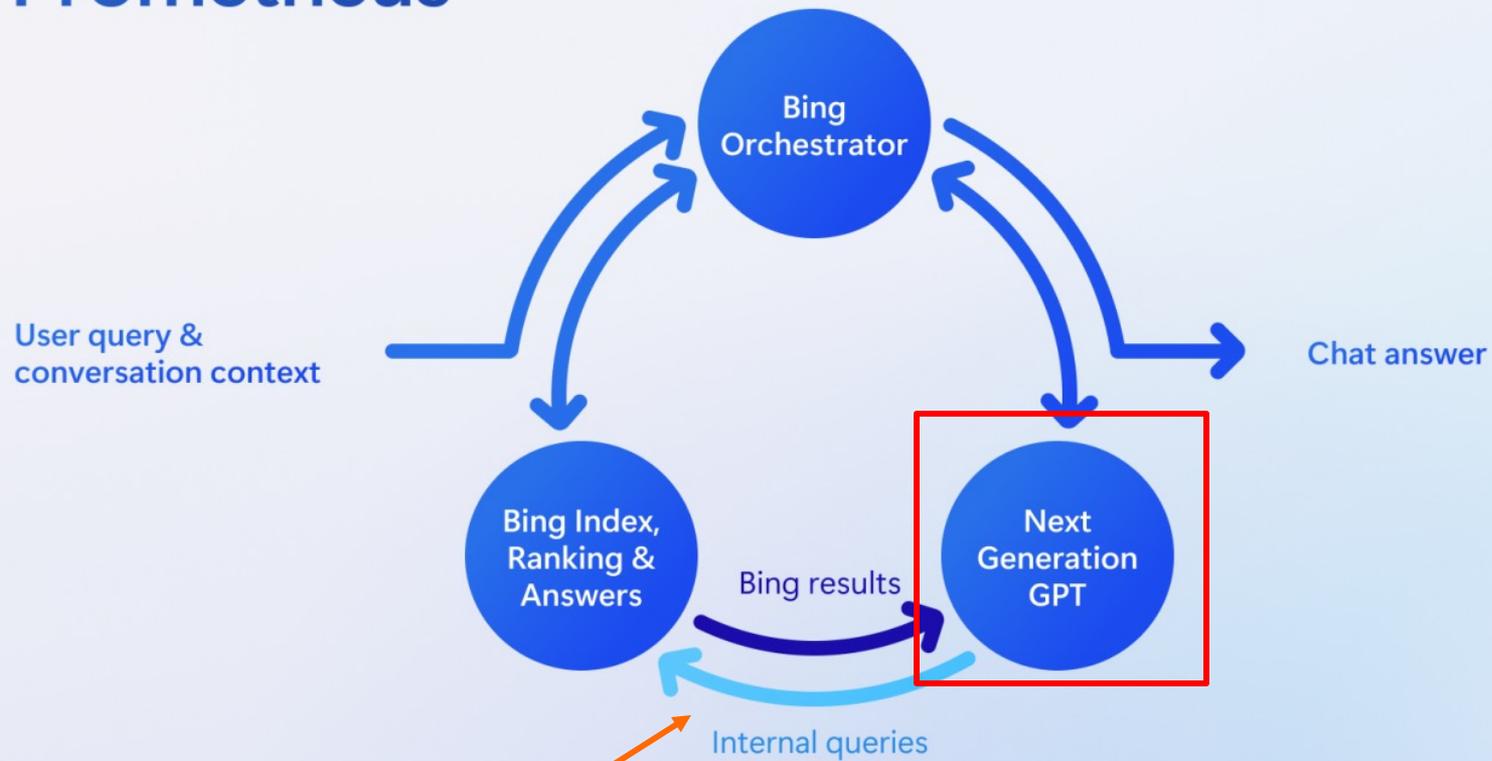
- IR-based methods directly retrieve and rank answers from the KBs considering the information conveyed in the questions. They consist of the following steps:
 1. Starting from the topic entity, the system first extracts a question-specific graph from KBs, ideally including all question-related entities and relations as nodes and edges.
 2. Next, the system encodes input questions into vectors representing reasoning instructions.
 3. A graph-based reasoning module conducts semantic matching via vector-based computation to propagate and then aggregate the information along the neighboring entities within the graph.
 4. An answer ranking module is utilized to rank the entities in the graph according to the reasoning status at the end of the reasoning phase. The top-ranked entities are predicted as the answers to the question.

CHALLENGES & SOLUTIONS



OPEN-DOMAIN QA (THE NEW BING)

Prometheus



Grounding

NEXT-GEN GPT

- Prompt-based Models: GPT-3, GPT-3.5, GPT-4
- User enters a prompt, system returns with a completion:

Prompt	Completion
Four score and seven years ago our	fathers brought forth on this continent, a new nation, conceived in Liberty, and dedicated to the proposition that all men are created equal. [...]
"Call me Ishmael."	"Some years ago—never mind how long precisely—having little or no money in my purse, and nothing particular to interest me on shore, I thought I would sail about a little and see the watery part of the world." [...]

- Model generates the remaining tokens given the prompt as a “context”.

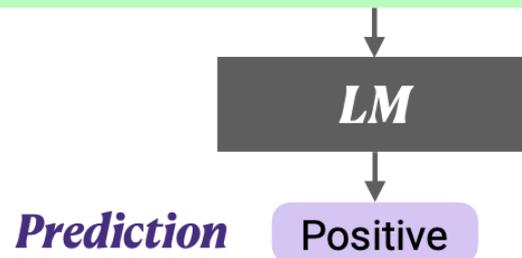
PROMPT-ENGINEERING

- Finding the best prompt for a task is a “witchcraft”.
- In-context learning by demonstrations:
 - Sentiment classification task

Demonstrations

Circulation revenue has increased by 5% in Finland.	\n	Positive
Panostaja did not disclose the purchase price.	\n	Neutral
Paying off the national debt will be extremely painful.	\n	Negative
The acquisition will have an immediate positive impact.	\n	_____

Test input

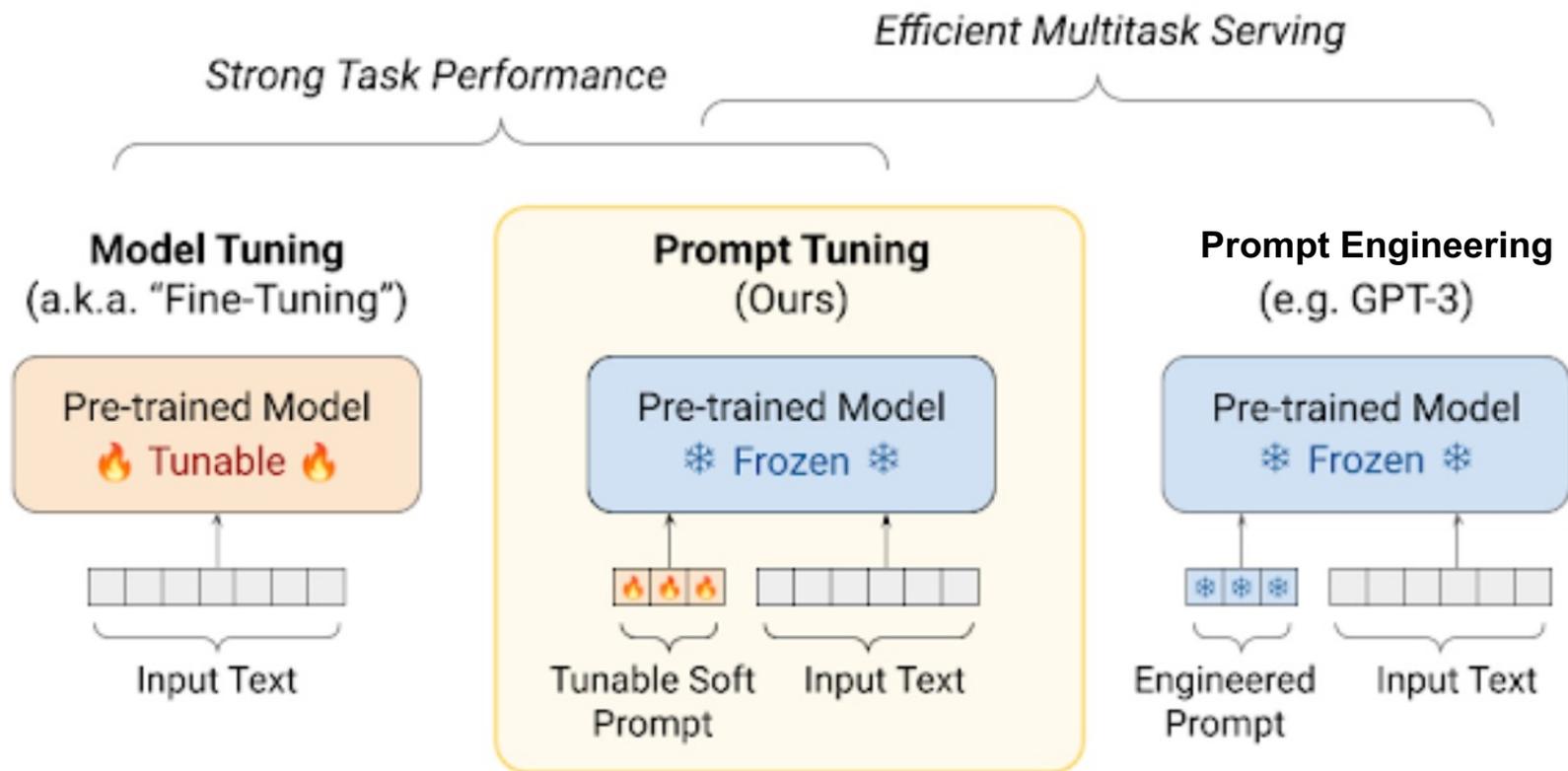


- Gold labels are not necessary!

QUIZ (PROMPT-ENGINEERING)

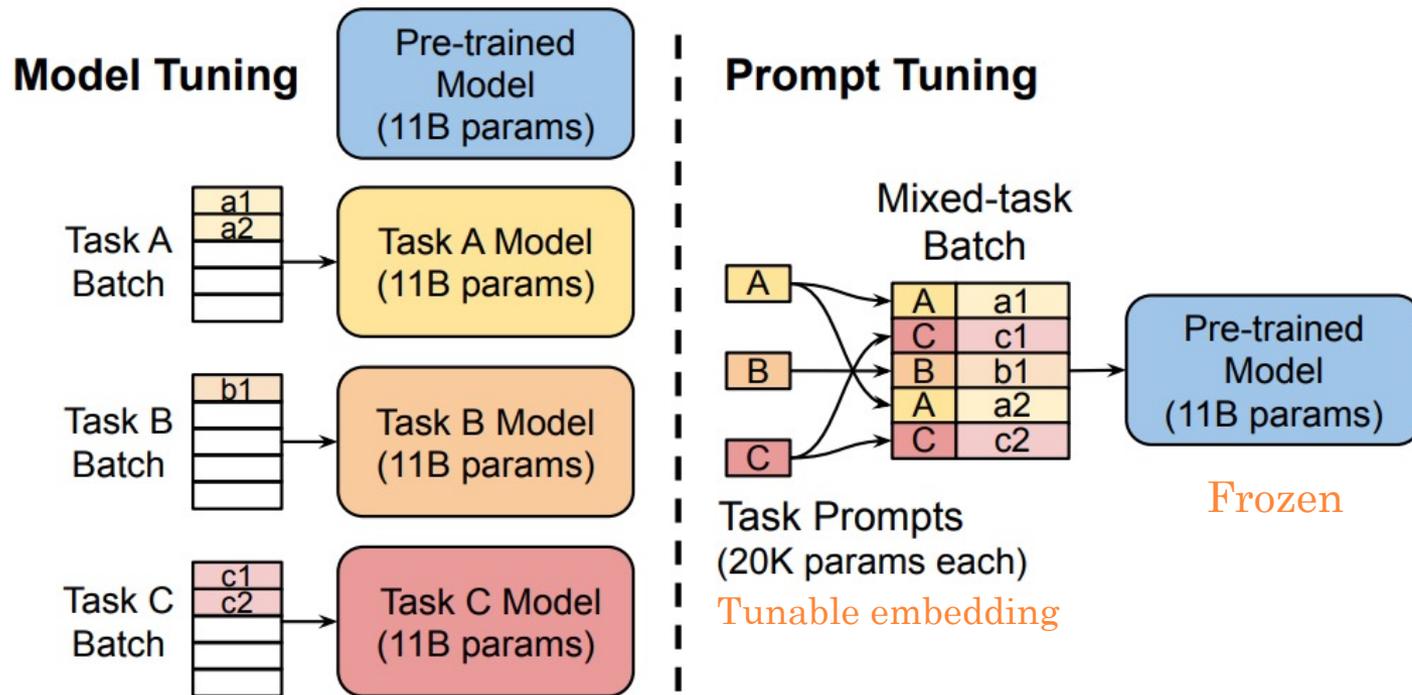
- Can you produce a prompt for doing English-to-Chinese (or English to any other language) translation on GPT-3, using few-shot in-context learning?

PROMPT-TUNING



PROMPT-TUNING

- Automatically generate “soft” prompts or a sequence of embeddings:



CHATGPT (INSTRUCTGPT)

Step 1

Collect demonstration data and train a supervised policy.

A prompt is sampled from our prompt dataset.



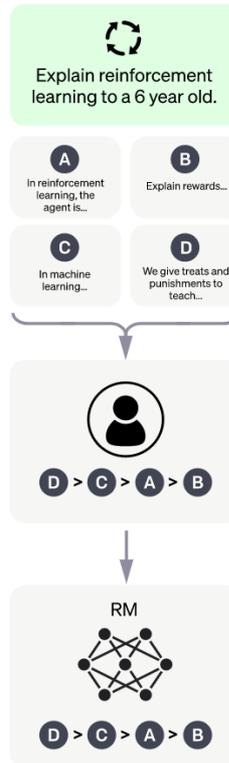
A labeler demonstrates the desired output behavior.

This data is used to fine-tune GPT-3.5 with supervised learning.

Step 2

Collect comparison data and train a reward model.

A prompt and several model outputs are sampled.



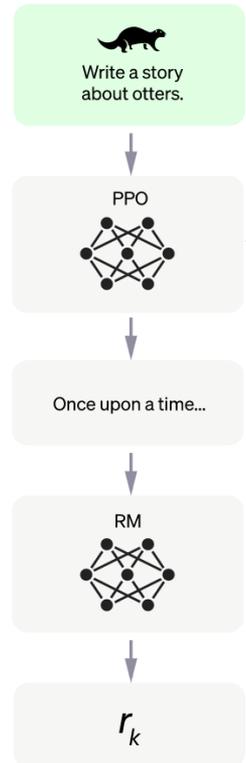
A labeler ranks the outputs from best to worst.

This data is used to train our reward model.

Step 3

Optimize a policy against the reward model using the PPO reinforcement learning algorithm.

A new prompt is sampled from the dataset.



The PPO model is initialized from the supervised policy.

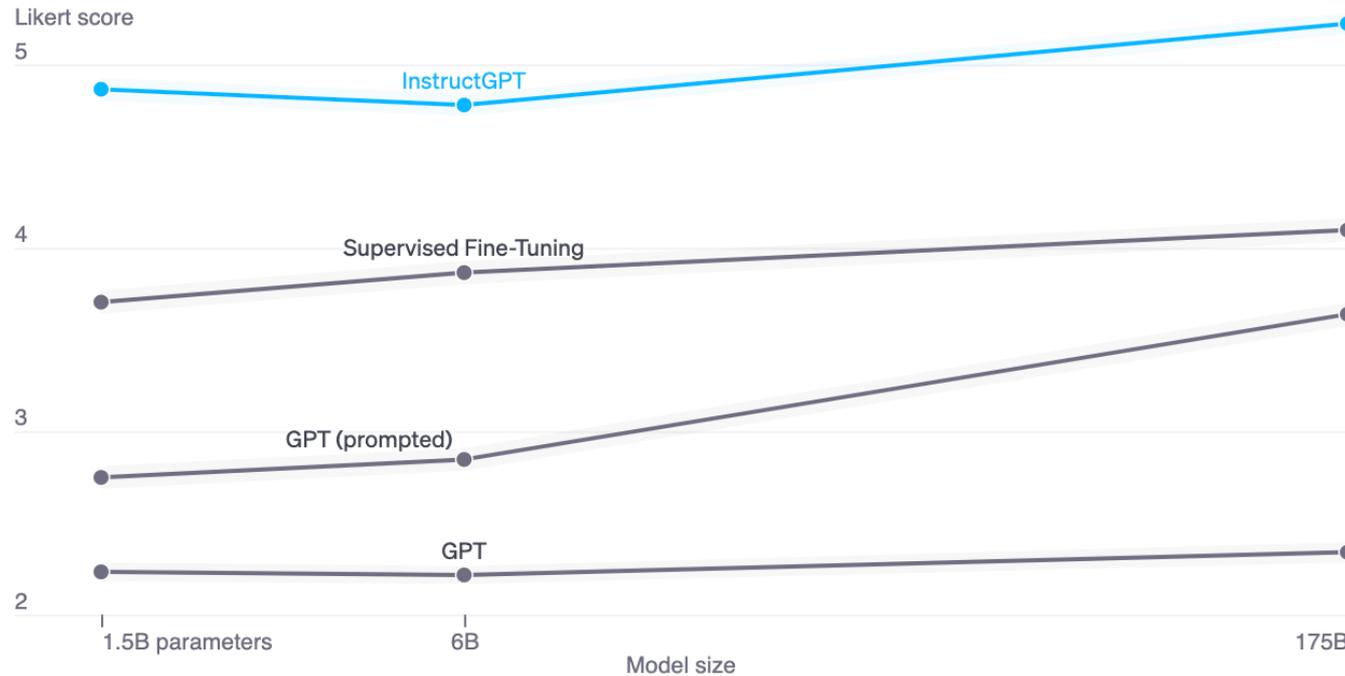
The policy generates an output.

The reward model calculates a reward for the output.

The reward is used to update the policy using PPO.

“Proximal Policy Optimization”

EVALUATIONS

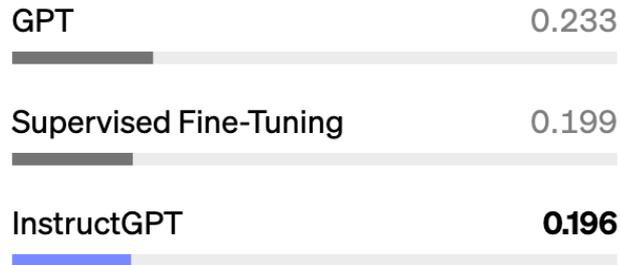


Quality ratings of model outputs on a 1–7 scale (y-axis), for various model sizes (x-axis), on prompts submitted to InstructGPT models on our API. InstructGPT outputs are given much higher scores by our labelers than outputs from GPT-3 with a few-shot prompt and without, as well as models fine-tuned with supervised learning. We find similar results for prompts submitted to GPT-3 models on the API.

EVALUATIONS

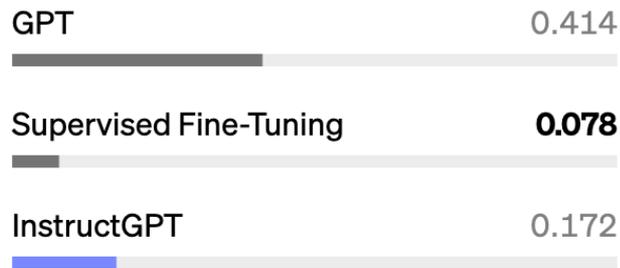
Dataset

RealToxicity



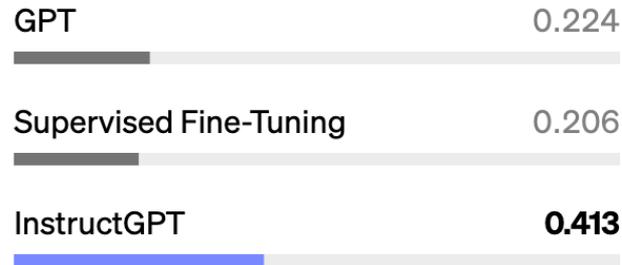
API Dataset

Hallucinations



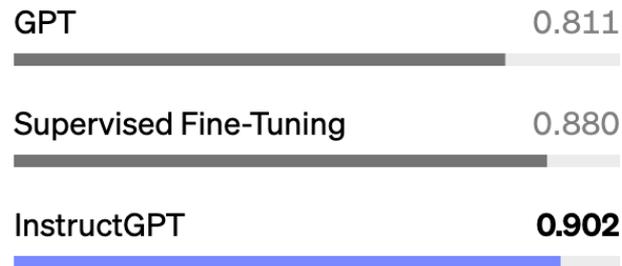
Dataset

TruthfulQA



API Dataset

Customer Assistant Appropriate

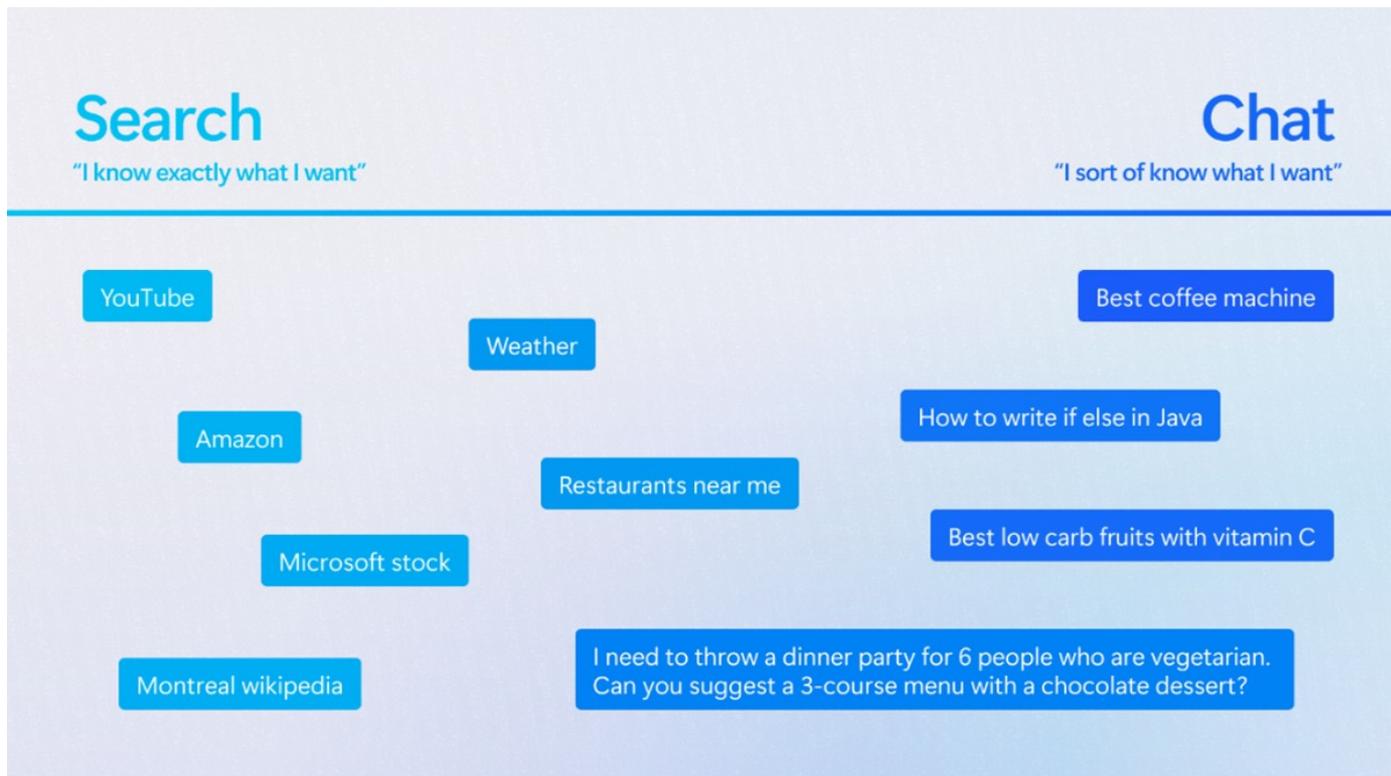


LIMITATIONS OF CHATGPT

- Produce nonsense text
 - there's no source of truth in reinforcement learning
- Excessive moderation
 - the question may contain no problem but the model refuses to answer
- Doesn't learn the gold answer
 - answer may not be as accurate as an expert would like
- Repetition of sentence or part of a sentence
 - bias toward long answers written by human experts
 - overlearning

UX DESIGN

- Search or Chat?
- Need to smoothly transition between Search and Chat modes based on the user's intent and preference.
- This is as important as the Prometheus framework.



I am planning a trip for our anniversary in September. What are some places we can go that are within a 3 hour flight from London Heathrow?

SEARCH CHAT WORK TRAVEL IMAGES VIDEOS MAPS MORE

About 76,400,000 results Any time Open links in new tab

The Best September Anniversary Travel Ideas - TripSavvy
https://www.tripsavvy.com/best-september-anniversary-travel-ideas-4173621
Web Jun 26, 2019 · Whether it's your fifth wedding anniversary, your 20th, or the anniversary of your first date, traveling together is a great way to celebrate it. And September is the best ...

Fall

Consider visiting more than one island when you go. Somewhere around the ...

Overwater Bungalows

Elizabeth Brownfield is a writer, editor, and researcher with over 15 years of ...

Other content from tripsavvy.com

United States - The Best September Anniversary Travel Ideas

Destinations - The Best September Anniversary Travel Ideas

Europe - The Best September Anniversary Travel Ideas

See more

Most Romantic Anniversary Trip Ideas - Travel + Leisure
https://www.travelandleisure.com/trip-ideas/amazing-anniversary-trip-ideas
Author: Alex Schechter

- 1. The French Riviera. Jetting off to the south of France to spend your days ...
2. Steamboat Springs, Colorado. If an active vacation is more your speed, head ...
3. Amsterdam, Netherlands. Sure, massive European capitals like London and ...
4. Los Cabos, Mexico. Fly south to Mexico's stunning beaches for a relaxing stay ...



See full list on travelandleisure.com

9 Anniversary Trip Ideas to Celebrate Any Milestone
https://www.weddingwire.com/wedding-ideas/anniversary-trip-ideas

- Napa, California. If you and your spouse appreciate a good glass of cabernet, chances are, ...
• The Bahamas. This tropical island just off the coast of Florida is a vacation hotspot for ...
• New York City. As the saying goes, "New York City is always a good idea," and this is ...
• Cabo San Lucas, Mexico. With low cases of COVID, Cabo San Lucas is a hot destination for ...

See full list on weddingwire.com

People also ask



Congratulations on your anniversary! 🎉 There are many places you can go that are within a 3 hour flight from London Heathrow. Here are some suggestions based on your preferences and the best destinations in Europe in September 4 5 6 :

- If you like beaches and sunshine, you can fly to Malaga in Spain 1, which has a pleasant climate, a lively culture, and a beautiful coastline. You can enjoy the sandy beaches, the historic monuments, and the delicious tapas. Malaga is also a great base to explore other nearby attractions, such as Granada, Seville,

Learn more: 1. flightconnections.com +3 more

? Thank you, that's very helpful.

How can I book a flight to Malaga?

What are some other places to visit in Europe?

Let's chat



- If you like mountains and lakes, you can fly to **Annecey** in France ¹, which is a charming alpine town near the Swiss border. Annecey is known as the "Venice of France" because of its canals and bridges. You can admire the stunning scenery, visit the medieval castle, and try some local cheese and wine. Annecey is also close to Geneva, Chamonix, and Lyon, if you want to see more of the region.
- If you like **art** and **history**, you can fly to **Florence** in Italy ², which is the birthplace of the Renaissance and a UNESCO World Heritage Site. Florence is a treasure trove of artistic and architectural masterpieces, such as the Duomo, the Uffizi Gallery, and the Ponte Vecchio. You can also explore the Tuscan countryside, taste the famous gelato, and shop for leather goods.
- If you like **culture** and **nightlife**, you can fly to **Berlin** in Germany ³, which is a vibrant and cosmopolitan city with a rich and diverse heritage. Berlin is a hub of creativity, innovation, and entertainment, with many museums, galleries, clubs, and festivals. You can also learn about the city's history, from the Brandenburg Gate to the Berlin Wall, and enjoy the multicultural cuisine and beer.

I hope this helps you plan your trip. Have a wonderful anniversary! 😊

Learn more: [1. flightconnections.com](#) [2. withinhours.com](#) [3. withinhours.com](#) [+3 more](#)

Can you recommend some hotels in Malaga? I'd like to stay in the center of the city.

✓ Searching for: **hotels in Malaga center**

✓ Generating answers for you...

Sure, I can recommend some hotels in Malaga center. Here are some options based on the ratings, reviews, and prices from Booking.com ¹:

- If you want a **luxurious** and **beachfront** hotel, you can stay at the **Gran Hotel Miramar GL** ¹, which is a 5-star hotel in a listed XX century building. It has a spa center, a seasonal pool, and beautiful rooms. The staff are very kind and helpful, and they take care of every detail. The hotel is 33 feet from La Malagueta Beach and close to the city center. The price per night is around \$300.
- If you want a **modern** and **stylish** hotel, you can stay at the **Hotel Brö-Adults Only** ¹ ², which is a 4-star hotel with an outdoor swimming pool, a garden, and a terrace. It has air-conditioned rooms, free

Learn more: [1. booking.com](#) [2. booking.com](#)

■ Stop Responding



Ask me anything...

0/2000



RESOURCES

Encoder-decoder seq2seq model explained: <https://medium.com/analytics-vidhya/encoder-decoder-seq2seq-models-clearly-explained-c34186fbf49b>

RNN explained: <https://www.analyticsvidhya.com/blog/2022/03/a-brief-overview-of-recurrent-neural-networks-rnn/>

Hochreiter, Sepp; Schmidhuber, Jürgen (1997-11-01). "Long Short-Term Memory". *Neural Computation*. 9 (8): 1735–1780.

Attention is all you need! <https://arxiv.org/abs/1706.03762>

Lan, Yunshi, et al. "A survey on complex knowledge base question answering: Methods, challenges and solutions." *arXiv preprint arXiv:2105.11644* (2021).

Michihiro Yasunaga, Hongyu Ren, Antoine Bosselut, Percy Liang, Jure Leskovec. QA-GNN: Reasoning with Language Models and Knowledge Graphs for Question Answering. NAACL 2021.

<https://www.microsoft.com/en-us/research/uploads/prod/2020/07/SIGIR-2020-Tutorial-Slides-recent-advances-in-conversational-information-retrieval.pdf>

Brown, Tom, et al. "Language models are few-shot learners." *Advances in neural information processing systems* 33 (2020): 1877-1901.

Lester, Brian, Rami Al-Rfou, and Noah Constant. "The power of scale for parameter-efficient prompt tuning." *EMNLP 2021*.

Li, Xiang Lisa, and Percy Liang. "Prefix-tuning: Optimizing continuous prompts for generation." *arXiv preprint arXiv:2101.00190* (2021).

Ouyang, Long, et al. "Training language models to follow instructions with human feedback." *Advances in Neural Information Processing Systems* 35 (2022): 27730-27744.

<https://pub.towardsai.net/chatgpt-how-does-it-work-internally-e0b3e23601a1>

<https://blogs.bing.com/search-quality-insights/february-2023/Building-the-New-Bing>