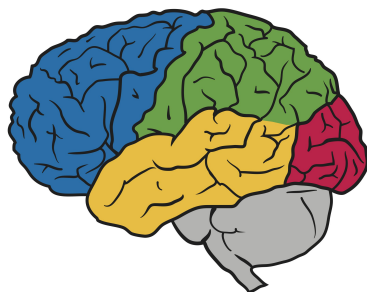# Answering Questions by Querying the Implicit Knowledge Base Inside T5

Colin Raffel
with Adam Roberts, Noam Shazeer, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J. Liu

The cabs ___ the same rates as those ___ by horse-drawn cabs and were ___ quite popular, ___ the Prince of Wales (the ___ King Edward VII) travelled in ___. The cabs quickly ___ known as "hummingbirds" for ___ noise made by their motors and their distinctive black and ___ livery. Passengers ___ ___ the interior fittings were ___ when compared to ___ cabs but there ___ some complaints ___ the ___ lighting made them too ___ to those outside ___.
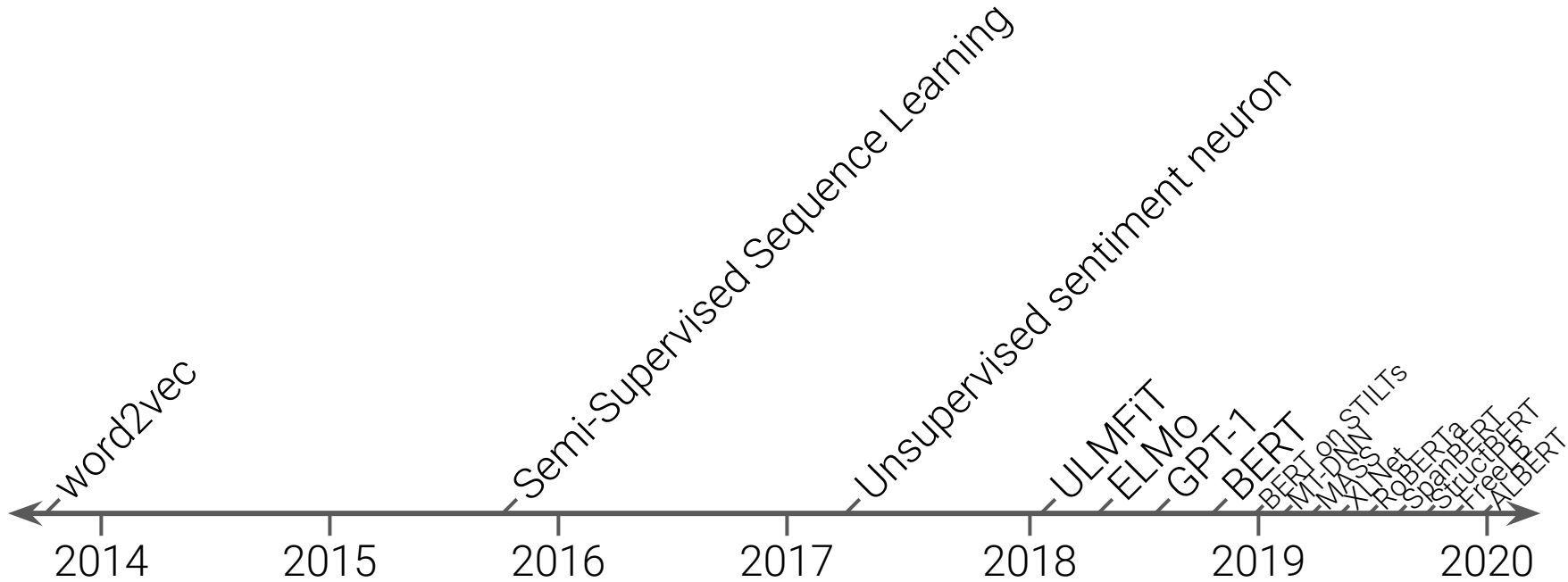
charged, used, initially, even, future, became, the, yellow, reported, that, luxurious, horse-drawn, were that, internal, conspicuous, cab

This movie is terrible! The acting is bad and I was bored the entire time. There was no plot and nothing interesting happened. I was really surprised since I had very high expectations. I want 103 minutes of my life back!

negative

word2vec

Semi-Supervised Sequence Learning

Unsupervised sentiment neuron

ULMFiT

ELMo

GPT-1

BERT

BERT on STILTs

MT-DNN

MASS

XLNet

RoBERTa

SpanBERT

StructBERT

FreeLB

ALBERT

2014   2015   2016   2017   2018   2019   2020

"translate English to German: That is good." → T5 → "Das ist gut."

"cola sentence: The course is jumping well." → T5 → "not acceptable"

"stsb sentence1: The rhino grazed on the grass. sentence2: A rhino is grazing in a field." → T5 → "3.8"

"summarize: state authorities dispatched emergency crews tuesday to survey the damage after an onslaught of severe weather in mississippi…" → T5 → "six people hospitalized after a storm in attala county."

Text-to-Text Transfer Transformer

Common Crawl Web Extracted Text

Menu

Lemon

Introduction

The lemon, Citrus Limon (l.) Osbeck, is a species of small evergreen tree in the flowering plant family rutaceae.
The tree's ellipsoidal yellow fruit is used for culinary and non-culinary purposes throughout the world, primarily for its juice, which has both culinary and cleaning uses.
The juice of the lemon is about 5% to 6% citric acid, with a ph of around 2.2, giving it a sour taste.

Article

The origin of the lemon is unknown, though lemons are thought to have first grown in Assam (a region in northeast India), northern Burma or China.
A genomic study of the lemon indicated it was a hybrid between bitter orange (sour orange) and citron.

Please enable JavaScript to use our site.

Home
Products
Shipping
Contact
FAQ

Dried Lemons, $3.59/pound

Organic dried lemons from our farm in California.
Lemons are harvested and sun-dried for maximum flavor.
Good in soups and on popcorn.

The lemon, Citrus Limon (l.) Osbeck, is a species of small evergreen tree in the flowering plant family rutaceae.
The tree's ellipsoidal yellow fruit is used for culinary and non-culinary purposes throughout the world, primarily for its juice, which has both culinary and cleaning uses.
The juice of the lemon is about 5% to 6% citric acid, with a ph of around 2.2, giving it a sour taste.

Lorem ipsum dolor sit amet, consectetur adipiscing elit.
Curabitur in tempus quam. In mollis et ante at consectetur.
Aliquam erat volutpat.
Donec at lacinia est.
Duis semper, magna tempor interdum suscipit, ante elit molestie urna, eget efficitur risus nunc ac elit.
Fusce quis blandit lectus.
Mauris at mauris a turpis tristique lacinia at nec ante.
Aenean in scelerisque tellus, a efficitur ipsum.
Integer justo enim, ornare vitae sem non, mollis fermentum lectus.
Mauris ultrices nisl at libero porta sodales in ac orci.

```
function Ball(r) {
    this.radius = r;
    this.area = pi * r ** 2;
    this.show = function(){
        drawCircle(r);
    }
}
```

Colossal Clean Crawled Corpus (C4)

Menu

Lemon

Introduction

The lemon, Citrus Limon (l.) Osbeck, is a species of small evergreen tree in the flowering plant family rutaceae.
The tree's ellipsoidal yellow fruit is used for culinary and non-culinary purposes throughout the world, primarily for its juice, which has both culinary and cleaning uses. The juice of the lemon is about 5% to 6% citric acid, with a ph of around 2.2, giving it a sour taste.

Article

The origin of the lemon is unknown, though lemons are thought to have first grown in Assam (a region in northeast India), northern Burma or China.
A genomic study of the lemon indicated it was a hybrid between bitter orange (sour orange) and citron.

Please enable JavaScript to use our site.

Home
Products
Shipping
Contact
FAQ

Dried Lemons, $3.59/pound

Organic dried lemons from our farm in California.
Lemons are harvested and sun-dried for maximum flavor.
Good in soups and on popcorn.

The lemon, Citrus Limon (l.) Osbeck, is a species of small evergreen tree in the flowering plant family rutaceae.
The tree's ellipsoidal yellow fruit is used for culinary and non-culinary purposes throughout the world, primarily for its juice, which has both culinary and cleaning uses. The juice of the lemon is about 5% to 6% citric acid, with a ph of around 2.2, giving it a sour taste.

Lorem ipsum dolor sit amet, consectetur adipiscing elit.
Curabitur in tempus quam. In mollis et ante at consectetur.
Aliquam erat volutpat.
Donec at lacinia est.
Duis semper, magna tempor interdum suscipit, ante elit molestie urna, eget efficitur risus nunc ac elit.
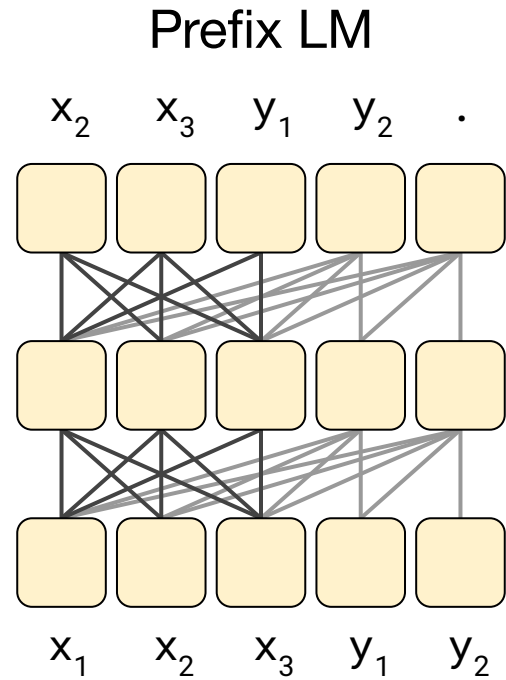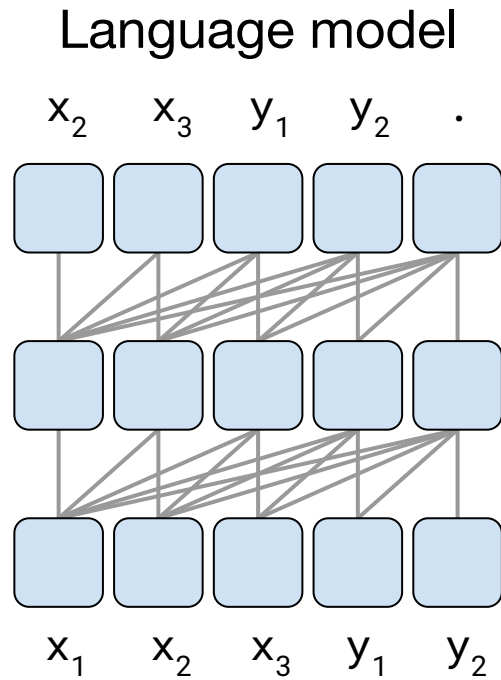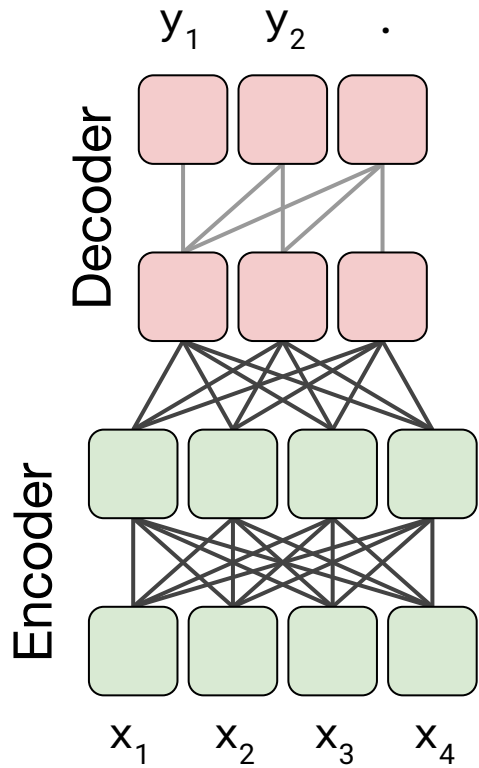Fusce quis blandit lectus.
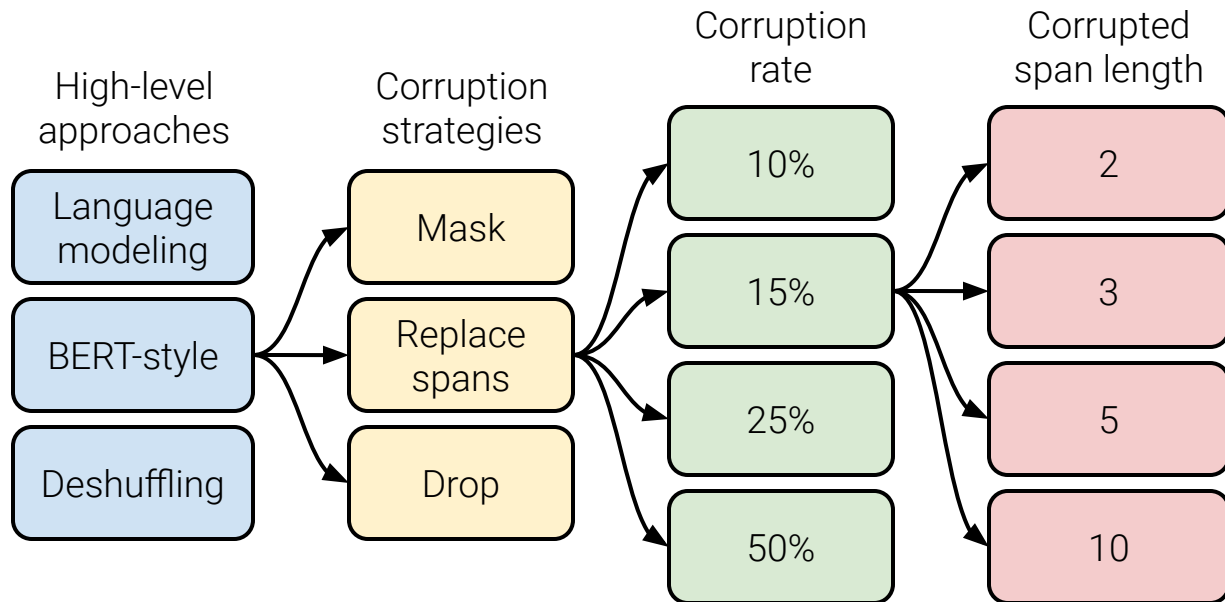Mauris at mauris a turpis tristique lacinia at nec ante.
Aenean in scelerisque tellus, a efficitur ipsum.
Integer justo enim, ornare vitae sem non, mollis fermentum lectus.
Mauris ultrices nisl at libero porta sodales in ac orci.

```
function Ball(r) {
    this.radius = r;
    this.area = pi * r ** 2;
    this.show = function(){
        drawCircle(r);
    }
}
```

Objective	Inputs	Targets
Prefix language modeling	Thank you for inviting	me to your party last week .
BERT-style	Thank you `<M>` `<M>` me to your party apple week .	*(original text)*
Deshuffling	party me for your to . last fun you inviting week Thank	*(original text)*
I.i.d. noise, mask tokens	Thank you `<M>` `<M>` me to your party `<M>` week .	*(original text)*
I.i.d. noise, replace spans	Thank you `<X>` me to your party `<Y>` week .	`<X>` for inviting `<Y>` last `<Z>`
I.i.d. noise, drop tokens	Thank you me to your party week .	for inviting last
Random spans	Thank you `<X>` to `<Y>` week .	`<X>` for inviting me `<Y>` your party last `<Z>`

Training loss

Dataset size
- Full dataset
- 2^{29}
- 2^{27}
- 2^{25}
- 2^{23}

Step × 1,000
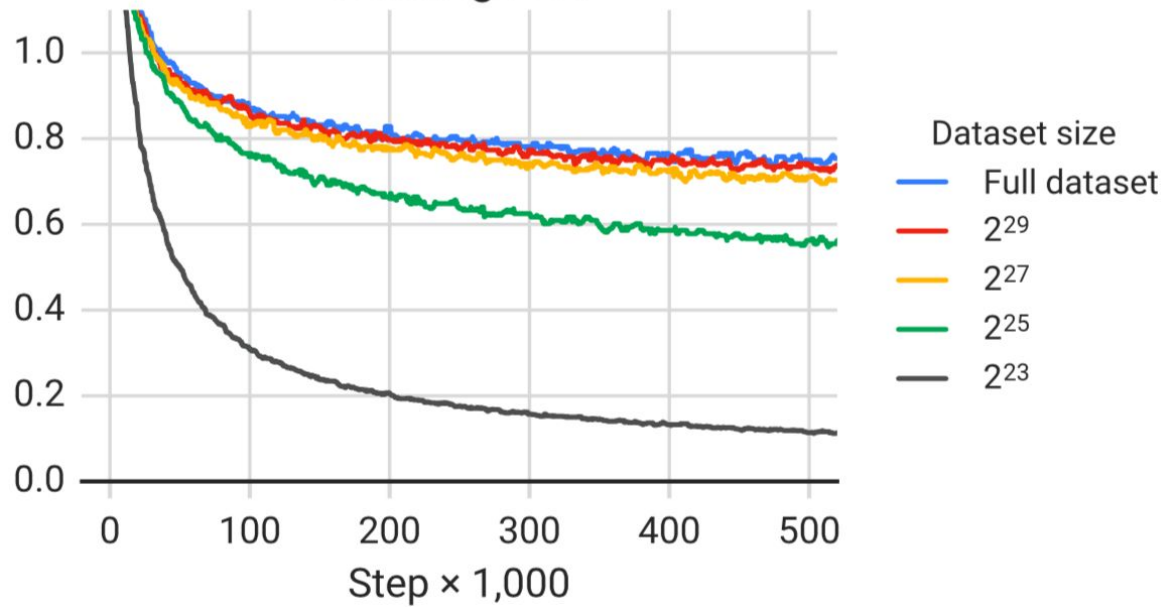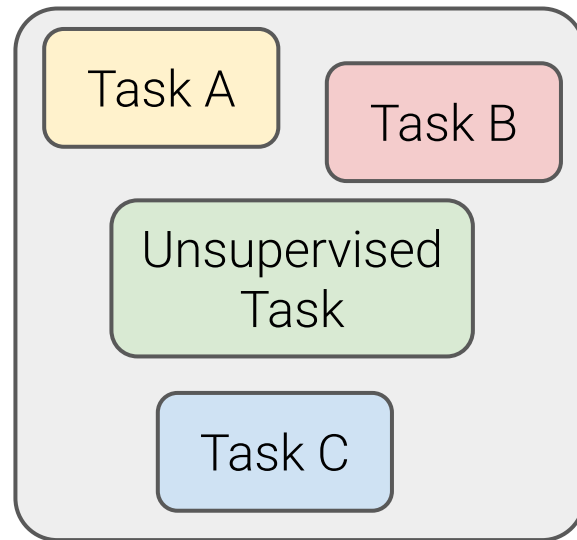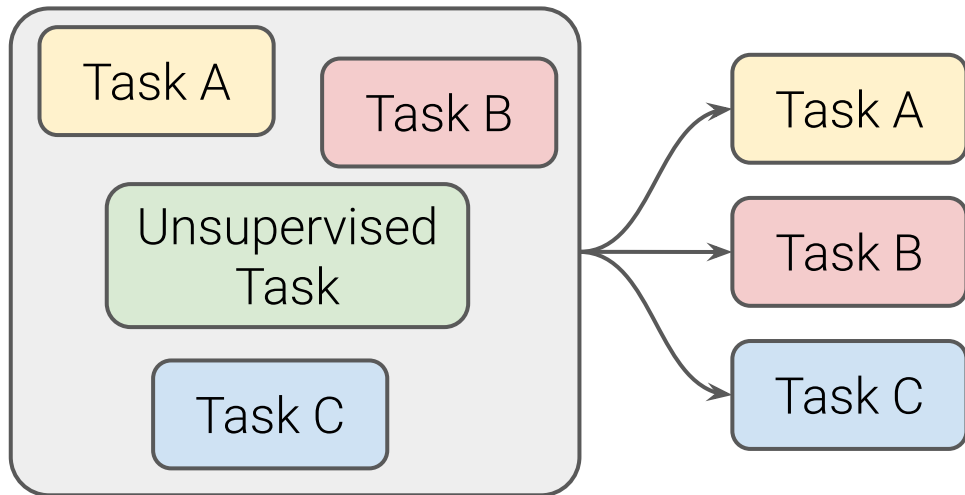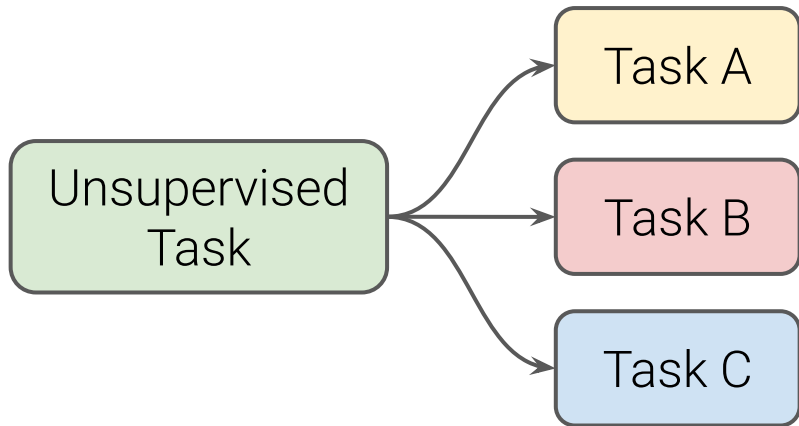
Model	GLUE Average	CNN/DM ROUGE-2-F	SQuAD EM	SuperGLUE Average	WMT EnDe BLEU	WMT EnFr BLEU	WMT EnRo BLEU
Previous best	89.4	20.30	90.1	84.6	**33.8**	**43.8**	**38.5**
T5-Small	77.4	19.56	87.24	63.3	26.7	36.0	26.8
T5-Base	82.7	20.34	92.08	76.2	30.9	41.2	28.0
T5-Large	86.4	20.68	93.79	82.3	32.0	41.5	28.1
T5-3B	88.5	21.02	94.95	86.4	31.8	42.6	28.2
T5-11B	**90.3**	**21.55**	**91.26**	**89.3**	32.1	43.4	28.1

▸ Copyright 2019 The T5 Authors

↳ 1 cell hidden

Fine-Tuning the Text-To-Text Transfer Transformer (T5) for Context-Free Trivia

Or: What does T5 know?

The following tutorial guides you through the process of fine-tuning a pre-trained T5 model, evaluating its accuracy, and using it for prediction, all on a free Google Cloud TPU **CO** **Open in Colab** *.*

http://tiny.cc/t5-colab

Reading Comprehension

Question

"What color is a lemon?"

Context

"The lemon tree's ellipsoidal yellow fruit is used for culinary and non-culinary purposes throughout the world, primarily for its juice, which has both culinary and cleaning uses. The pulp and rind are also used in cooking and baking."

Model

yellow

Open-Domain Question Answering

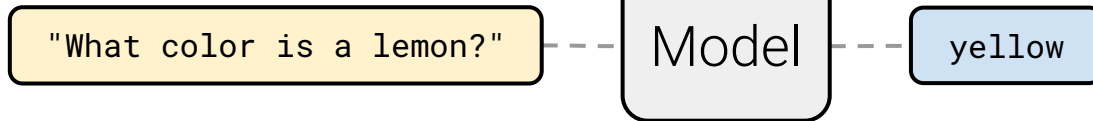*Question*

"What color is a lemon?"

Database

"The lemon tree's ellipsoidal yellow fruit is used for culinary and non-culinary purposes throughout the world, primarily for its juice, which has both culinary and cleaning uses. The pulp and rind are also used in cooking and baking."

Model

yellow

Closed-Book Question Answering

Question

"What color is a lemon?" - - - Model - - - yellow

President Franklin <M> born <M> January 1882.

D. Roosevelt was <M> in

Lily couldn't <M>. The waitress had brought the largest <M> of chocolate cake <M> seen.

believe her eyes <M> piece <M> she had ever

Our <M> hand-picked and sun-dried <M> orchard in Georgia.

peaches are <M> at our

T5

President Franklin D. Roosevelt was born in January 1882.

Pre-training

Fine-tuning

When was Franklin D. Roosevelt born?

T5

1882

	NQ	WQ	TQA
Open-domain SoTA	44.5	45.5	68.0
Closed-book SoTA	29.9	41.5	71.2

	NQ	WQ	TQA
Open-domain SoTA	44.5	45.5	68.0
Closed-book SoTA	29.9	41.5	71.2
T5-Base	27.0	29.1	29.1
T5-Large	29.8	32.2	35.9
T5-3B	32.1	34.9	43.4
T5-11B	34.5	37.4	50.1

<M> (born 1957) is a Spanish librarian who has been the director of the National Library of Spain since February 2013.
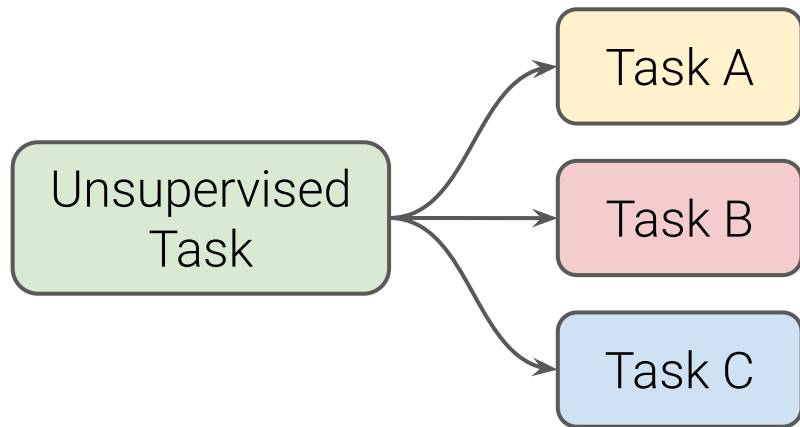
T5

Ana Santos Aramburo

TriviaQA

- Salient Span Masking
- Span Corruption
- Baseline

Validation EM
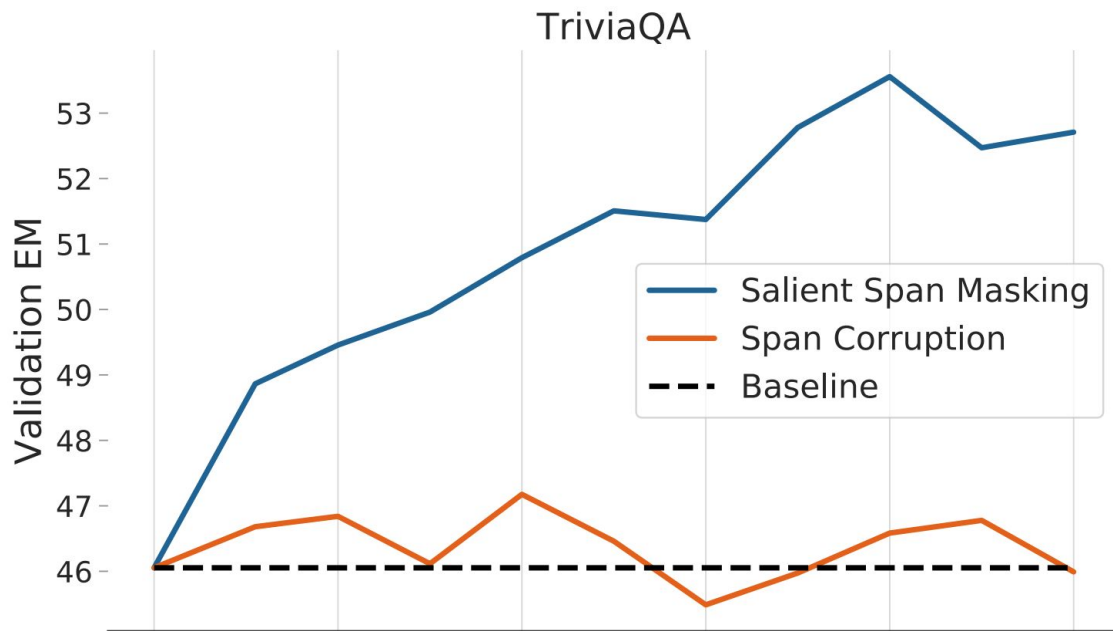
SSM data from "REALM: Retrieval-Augmented Language Model Pre-Training" by Guu et al.

	NQ	WQ	TQA
Open-domain SoTA	44.5	45.5	68.0
Closed-book SoTA	29.9	41.5	71.2
T5-Base	27.0	29.1	29.1
T5-Large	29.8	32.2	35.9
T5-3B	32.1	34.9	43.4
T5-11B	34.5	37.4	50.1
T5-11B + SSM	36.6	44.7	60.5
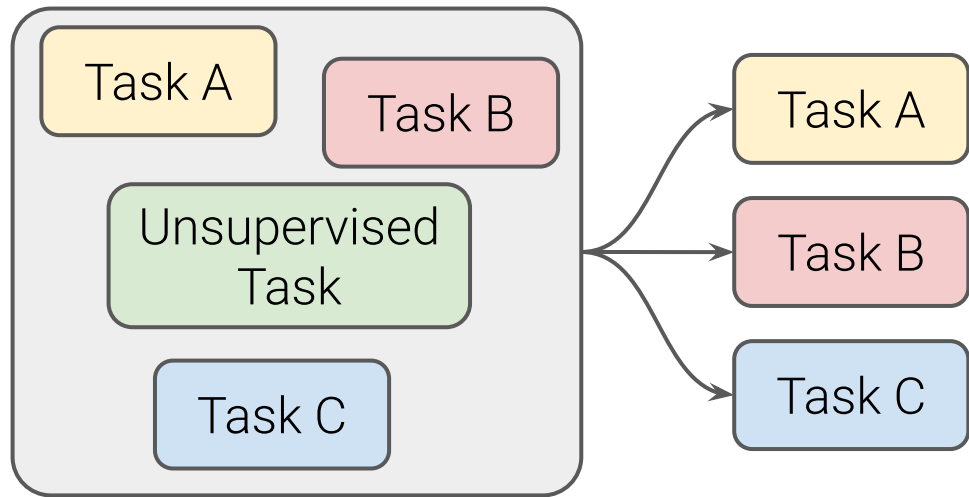
	NQ	WQ	TQA
Open-domain SoTA	44.5	45.5	68.0
Closed-book SoTA	29.9	41.5	71.2
T5-Base	27.0	29.1	29.1
T5-Large	29.8	32.2	35.9
T5-3B	32.1	34.9	43.4
T5-11B	34.5	37.4	50.1
T5-11B + SSM	36.6	44.7	60.5
T5.1.1-Base	26.8	28.8	30.6
T5.1.1-Large	28.9	30.8	37.2
T5.1.1-XL	32.2	33.8	45.1
T5.1.1-XXL	34.2	37.4	52.5
T5.1.1-XXL + SSM	37.9	43.5	61.6

Category	Question	Target(s)	T5 Prediction
True Negative	what does the ghost of christmas present sprinkle from his torch	little warmth, warmth	confetti
Phrasing Mismatch	who plays red on orange is new black	kate mulgrew	katherine kiernan maria mulgrew
Incomplete Annotation	where does the us launch space shuttles from	florida	kennedy lc39b
Unanswerable	who is the secretary of state for northern ireland	karen bradley	james brokenshire

	Category	Question	Target(s)	T5 Prediction
❌	True Negative	what does the ghost of christmas present sprinkle from his torch	little warmth, warmth	confetti
✓	Phrasing Mismatch	who plays red on orange is new black	kate mulgrew	katherine kiernan maria mulgrew
✓	Incomplete Annotation	where does the us launch space shuttles from	florida	kennedy lc39b
	Unanswerable	who is the secretary of state for northern ireland	karen bradley	james brokenshire

	12.5%	25%	37.5%	50%	62.5%

❌ True Negative

✔️ Phrasing mismatch

✔️ Incomplete annotation

🗑️ Unanswerable

Exact Match: 36.6 → 57.8%!

Cons	*Pros*

Cons	*Pros*
- Giant model (11 billion parameters)	- Actually smaller (?) than indices used for retrieval

Cons	*Pros*
- Giant model (11 billion parameters) - Expensive to run inference	- Actually smaller (?) than indices used for retrieval - Doesn't need to do retrieval or process context to produce answer

Cons	*Pros*
- Giant model (11 billion parameters)	- Actually smaller (?) than indices used for retrieval
- Expensive to run inference	- Doesn't need to do retrieval or process context to produce answer
- Can generate anything	- Can generate anything

Cons	*Pros*
- Giant model (11 billion parameters)	- Actually smaller (?) than indices used for retrieval
- Expensive to run inference	- Doesn't need to do retrieval or process context to produce answer
- Can generate anything	- Can generate anything
- Doesn't provide evidence	- Learns from giant, unstructured, noisy web text

Cons	*Pros*
- Giant model (11 billion parameters)	- Actually smaller (?) than indices used for retrieval
- Expensive to run inference	- Doesn't need to do retrieval or process context to produce answer
- Can generate anything	- Can generate anything
- Doesn't provide evidence	- Learns from giant, unstructured, noisy web text
- Not obvious how to add or remove facts	- Clear avenues for future work 😛

Cons	*Pros*
- Giant model (11 billion parameters)	- Actually smaller (?) than indices used for retrieval
- Expensive to run inference	- Doesn't need to do retrieval or process context to produce answer
- Can generate anything	- Can generate anything
- Doesn't provide evidence	- Learns from giant, unstructured, noisy web text
- Not obvious how to add or remove facts	- Clear avenues for future work 😛
- Not state-of-the-art	- Extremely simple approach that works surprisingly well

Cons	*Pros*
- Giant model (11 billion parameters)	- Actually smaller (?) than indices used for retrieval
- Expensive to run inference	- Doesn't need to do retrieval or process context to produce answer
- Can generate anything	- Can generate anything
- Doesn't provide evidence	- Learns from giant, unstructured, noisy web text
- Not obvious how to add or remove facts	- Clear avenues for future work 😛
- Not state-of-the-art	- Extremely simple approach that works surprisingly well
- Closed-book question answering is hard to evaluate	- Tests implicit knowledge with an important real-world problem

Adam Roberts

Noam Shazeer

Katherine Lee

Sharan Narang

Michael Matena

Yanqi Zhou

Wei Li

Peter J. Liu

Questions?