

Enabling Language Models to Fill in the Blanks



Chris Donahue



Mina Lee



Percy Liang



Paper <https://arxiv.org/abs/2005.05339>
Code <https://github.com/chrisdonahue/ilm>
Demo <https://chrisdonahue.com/ilm>

Why filling in the blanks?

Feedback on draft



From Mina Lee <minalee@cs.stanford.edu> ▾

Cc Bcc

To Chris Donahue (cs.stanford.edu) ✕

Feedback on draft

Hi Chris,

Thanks for updating the draft.

Can you revert the wording of the task definition?

Editing and revising

Why filling in the blanks?

Feedback on draft



From Mina Lee <minalee@cs.stanford.edu> ▾

Cc Bcc

To Chris Donahue (cs.stanford.edu) ✕

Feedback on draft

Hi Chris,

Thanks for updating the draft. The modifications look

Can you revert the wording of the task definition?

Editing and revising

Why filling in the blanks?

Feedback on draft



From Mina Lee <minalee@cs.stanford.edu> ▾

Cc Bcc

To Chris Donahue (cs.stanford.edu) ✕

Feedback on draft

Hi Chris,

Thanks for updating the draft. The modifications look great to me.

Can you revert the wording of the task definition?

Editing and revising

Why filling in the blanks?

Feedback on draft



From Mina Lee <minalee@cs.stanford.edu> ▾

Cc Bcc

To Chris Donahue (cs.stanford.edu) ✕

Feedback on draft

Hi Chris,

Thanks for updating the draft. The modifications look good with one exception.

Can you revert the wording of the task definition?

Editing and revising

Why filling in the blanks?

The image shows a screenshot of a Google Docs document. At the top, the document title is "Masterpiece". Below the title are icons for a star, a folder, and a cloud. To the right, there is a "Share" button and a profile picture of a woman. The menu bar includes "File", "Edit", "View", "Insert", "Format", "Tools", "Add-ons", and "Help". The status bar shows "Last edit was seconds a...". The toolbar contains icons for undo, redo, print, text color, background color, zoom (100%), text style (Normal text), font (Arial), font size (11), bold (B), italic (I), underline (U), text color (A), and a pencil icon. The main text area contains the sentence: "We were lost in the dark forest. Suddenly,".

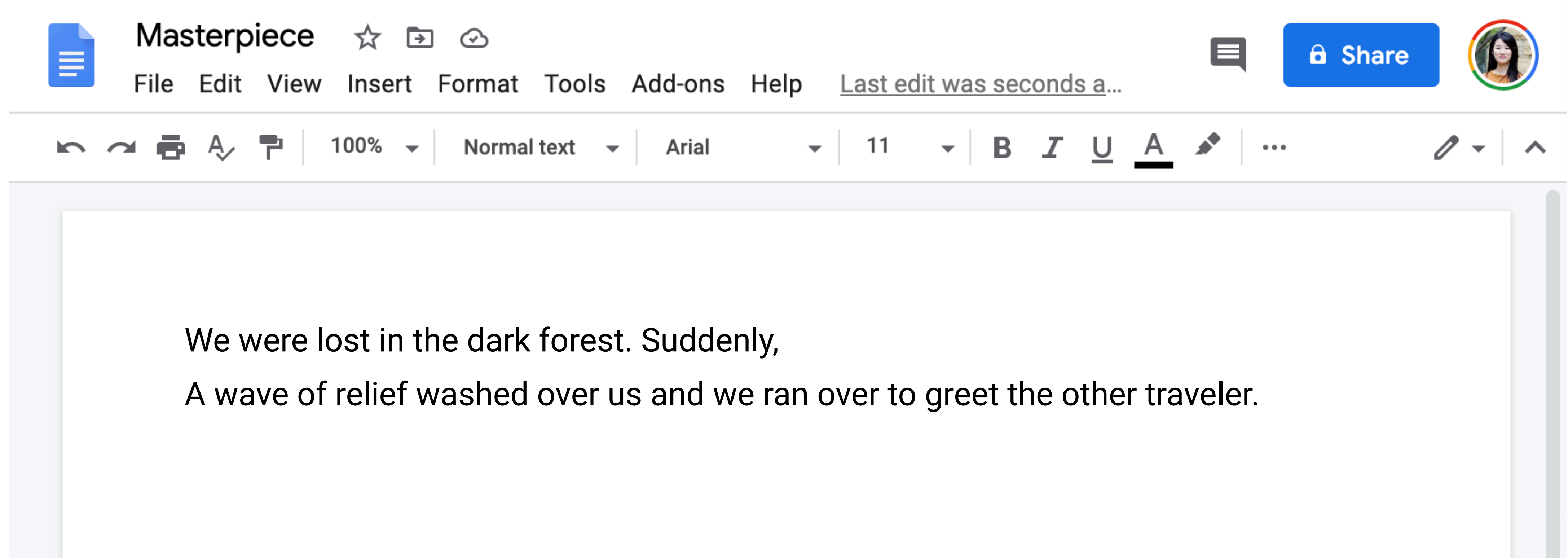
Connecting ideas

Why filling in the blanks?

A screenshot of a Google Docs interface. At the top, the document title is "Masterpiece" with icons for star, folder, and cloud. Below the title is a menu bar with "File", "Edit", "View", "Insert", "Format", "Tools", "Add-ons", and "Help". To the right of the menu bar is a "Share" button and a user profile picture. Below the menu bar is a toolbar with icons for undo, redo, print, text color, background color, zoom (100%), text style (Normal text), font (Arial), font size (11), bold (B), italic (I), underline (U), text color (A), and a pencil icon. The main editing area contains a single line of text: "We were lost in the dark forest. Suddenly, a bear emerged from the trees!".

Connecting ideas

Why filling in the blanks?

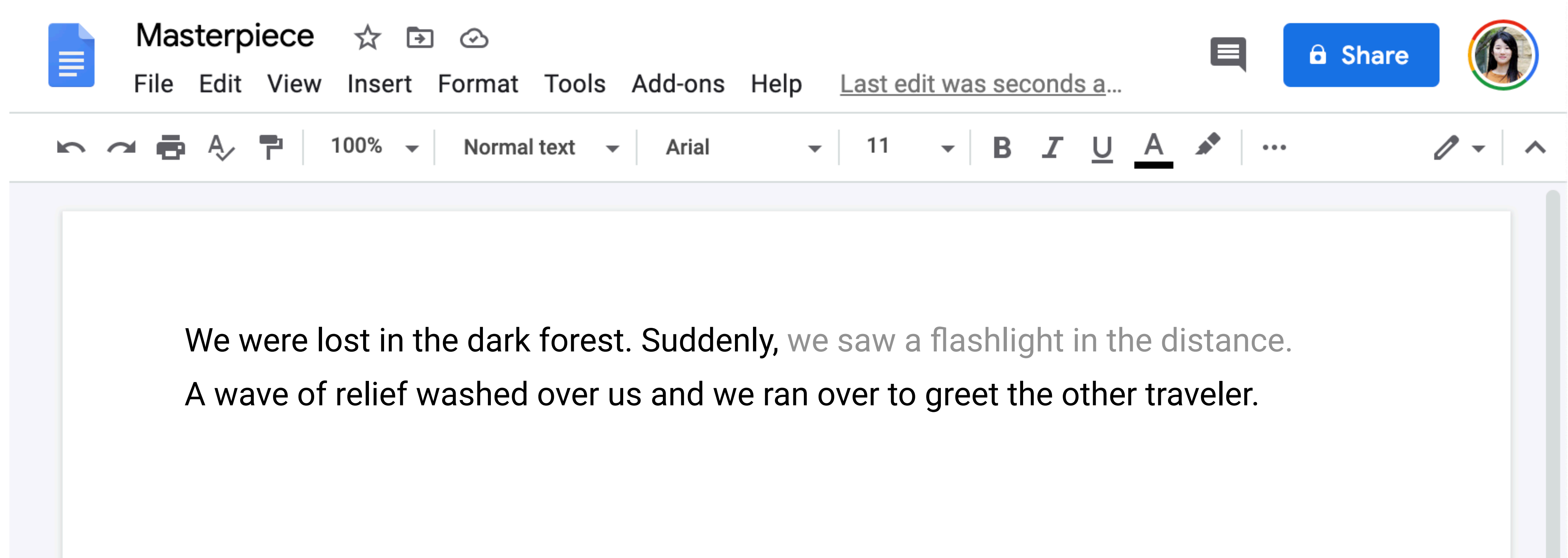


The image shows a screenshot of a Google Docs document. At the top, the document title is "Masterpiece". Below the title are icons for a star, a folder, and a cloud. To the right of these icons is a "Share" button with a lock icon and a profile picture of a woman. Below the title bar is a menu bar with options: File, Edit, View, Insert, Format, Tools, Add-ons, and Help. To the right of the menu bar is a status bar that says "Last edit was seconds a...". Below the menu bar is a toolbar with various editing tools: undo, redo, print, check for spelling, insert link, zoom (100%), text color (Normal text), font face (Arial), font size (11), bold (B), italic (I), underline (U), text color (A), and a pencil icon. The main content area of the document contains the following text:

We were lost in the dark forest. Suddenly,
A wave of relief washed over us and we ran over to greet the other traveler.

Connecting ideas

Why filling in the blanks?



The image shows a screenshot of a Google Docs document. At the top, the document title is "Masterpiece". Below the title are icons for a star, a folder, and a cloud. To the right of these icons is a "Share" button with a lock icon and a profile picture of a woman. Below the title bar is a menu bar with options: File, Edit, View, Insert, Format, Tools, Add-ons, and Help. To the right of the menu bar is a status bar that says "Last edit was seconds a...". Below the menu bar is a toolbar with various editing tools: undo, redo, print, check for spelling, insert link, zoom (100%), text color (Normal text), font face (Arial), font size (11), bold (B), italic (I), underline (U), text color (A), and a pencil icon. The main content area of the document contains the following text:

We were lost in the dark forest. Suddenly, we saw a flashlight in the distance.
A wave of relief washed over us and we ran over to greet the other traveler.

Connecting ideas

Text infilling

She ate [blank] for [blank].



She ate leftover pasta for lunch.

Input

Output

Given incomplete text with [blank]s, predict complete text



Arbitrary number of blanks

Variable length spans (e.g. word, sentence, paragraph)

Previous work on text infilling

She ate [blank] for [blank].



She ate leftover pasta for lunch.

Input

Output

General-purpose models

GPT-3 (Brown et al., 2020): Cannot consider future context

Previous work on text infilling

She ate [mask] [mask] for [mask].



She ate leftover pasta for lunch.

Input

Output

General-purpose models

GPT-3 (Brown et al., 2020): Cannot consider future context

BERT (Devlin et al., 2019): Must know **exact number of tokens**

Previous work on text infilling

She ate [blank] for [blank].



She ate leftover pasta for lunch.

Input

Output

General-purpose models

GPT-3 (Brown et al., 2020): Cannot consider future context

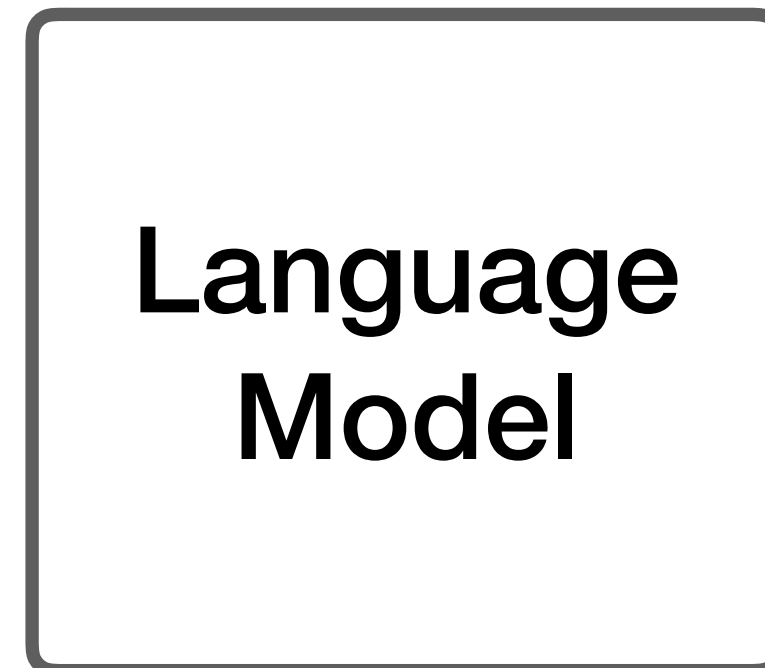
BERT (Devlin et al., 2019): Must know **exact number of tokens**

Task-specific models

SA (Zhu et al., 2019): Cannot leverage pre-trained language models

Our Idea: Infilling by Language Modeling (ILM)

1. Download your favorite language model (LM)



Our Idea: Infilling by Language Modeling (ILM)

1. Download your favorite language model (LM)
2. Fine-tune the model on infilling examples

Language
Model

[blank] for [blank]. [sep] leftover pasta [answer] lunch [answer]

Our Idea: Infilling by Language Modeling (ILM)

1. Manufacture infilling examples

Training time

Our Idea: Infilling by Language Modeling (ILM)

1. Manufacture infilling examples

Training time

She ate leftover pasta for lunch.

Data

Our Idea: Infilling by Language Modeling (ILM)

1. Manufacture infilling examples

Training time

She ate leftover pasta for lunch.

Data

She ate [blank] for [blank].

Input

Our Idea: Infilling by Language Modeling (ILM)

1. Manufacture infilling examples

Training time

She ate leftover pasta for lunch.

Data

She ate [blank] for [blank].

leftover pasta [answer] lunch [answer]

Input

Target

Our Idea: Infilling by Language Modeling (ILM)

Training time

1. Manufacture infilling examples

She ate leftover pasta for lunch.

Data

She ate [blank] for [blank]. [sep] leftover pasta [answer] lunch [answer]

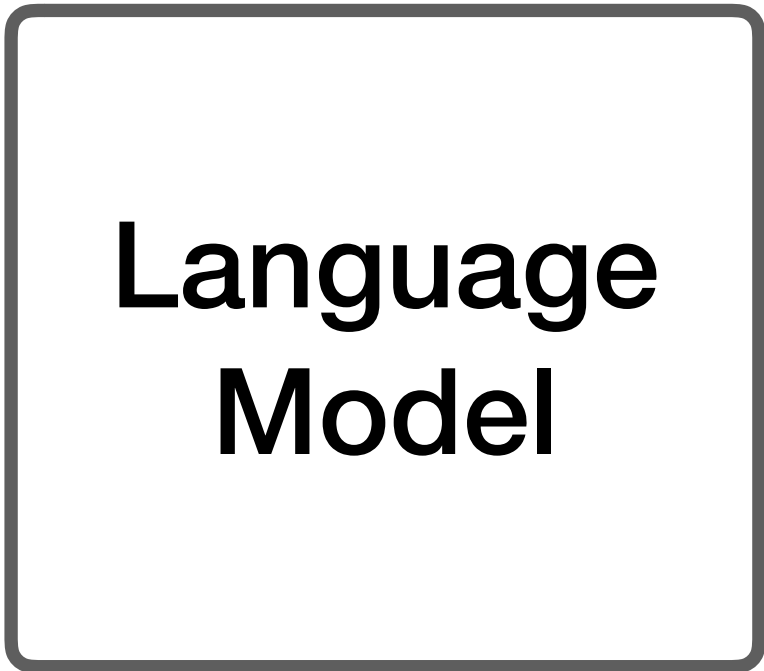
New data

Our Idea: Infilling by Language Modeling (ILM)

2. Download pre-trained left-to-right LM

Training time

She ate [blank] for [blank]. [sep] leftover pasta [answer] lunch [answer]

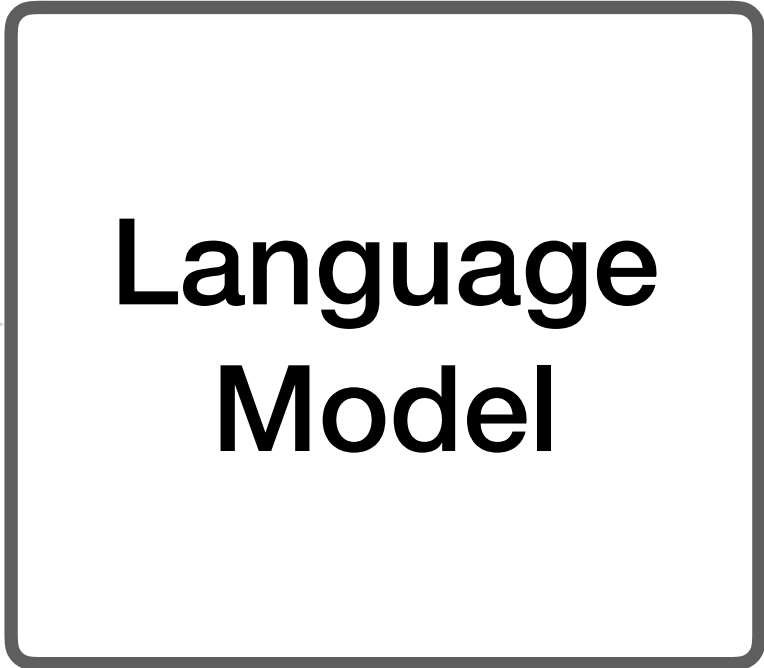


Our Idea: Infilling by Language Modeling (ILM)

3. Fine-tune LM on infilling examples

Training time

She ate [blank] for [blank]. [sep] leftover pasta [answer] lunch [answer]



Our Idea: Infilling by Language Modeling (ILM)

3. Fine-tune LM on infilling examples

Training time

Language
Model

[blank] for [blank]. [sep] leftover pasta [answer] lunch [answer]

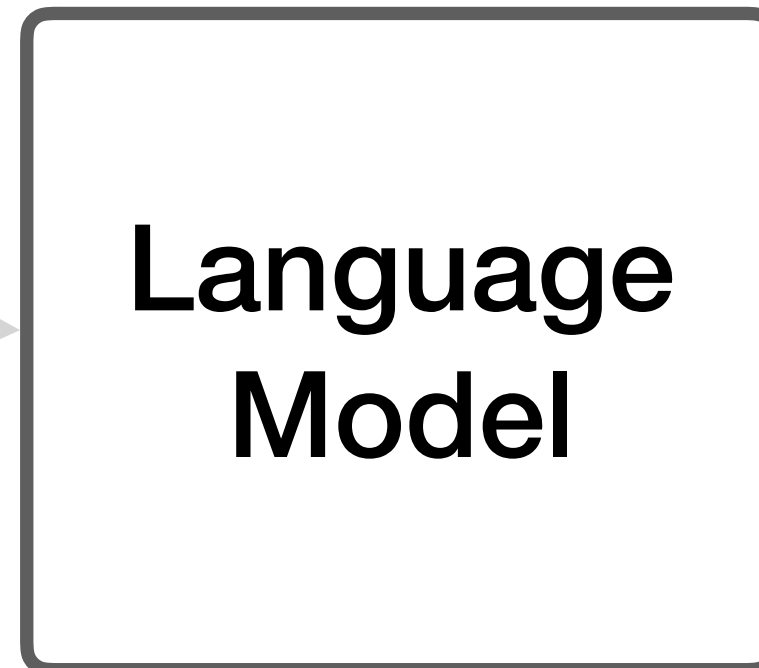
Our Idea: Infilling by Language Modeling (ILM)

Use fine-tuned LM to infill

Test time

He drinks [blank] after [blank]. [sep]

Input



Our Idea: Infilling by Language Modeling (ILM)

Use fine-tuned LM to infill

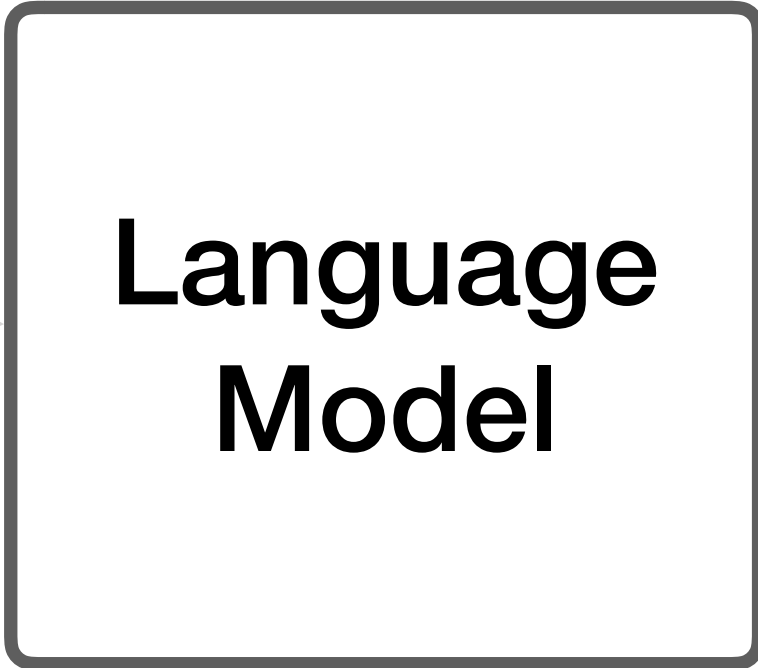
Test time

He drinks [blank] after [blank]. [sep]

water [answer] running [answer]

Input

Target



Our Idea: Infilling by Language Modeling (ILM)

Use fine-tuned LM to infill

Test time

He drinks [blank] after [blank]. [sep]

water [answer] running [answer]

Input

Target

He drinks water after running.

Output

Experimental setup

Data	Stories (Mostafazadeh et al., 2016), Abstracts, Lyrics
Metric	Score, Perplexity
Model	BERT, SA (Zhu et al., 2019), LM, ILM (ours)

- 1. Human evaluation**
- 2. Quantitative evaluation**

1. Human evaluation: Turing test

Identify one of the five sentences generated by **machine**.

Patty was excited about having her friends over.

She had been working hard preparing the food.

She also had the place looking spotless.

All of her friends arrived and were seated at the table.

Patty had a great time with her friends.

1. Human evaluation: Turing test

Identify one of the five sentences generated by **machine**.

Patty was excited about having her friends over.

She had been working hard preparing the food.

She also had the place looking spotless.

All of her friends arrived and were seated at the table.

Patty had a great time with her friends.

1. Human evaluation: Turing test

Identify one of the five sentences generated by **machine**.

Patty was excited about having her friends over.
She had been working hard preparing the food.

[blank]

All of her friends arrived and were seated at the table.
Patty had a great time with her friends.

1. Human evaluation: Turing test

Identify one of the five sentences generated by **machine**.

Patty was excited about having her friends over.
She had been working hard preparing the food.

[blank]

All of her friends arrived and were seated at the table.

Patty had a great time with her

ILM

Patty knew her friends wanted pizza.

1. Human evaluation: Turing test

Identify one of the five sentences generated by **machine**.

Patty was excited about having
She had been working hard pre
[blank]
All of her friends arrived and w
Patty had a great time with her

BERT	20%	favoritea ", Mary brightly said.
SA	29%	She wasn't sure she had to go to the store.
LM	41%	She went to check the tv.
ILM	45%	Patty knew her friends wanted pizza.

2. Quantitative evaluation

	Stories	Abstracts	Lyrics
LM	18.3	27.9	27.7
ILM	15.6	22.4	22.6

Perplexity on the sentence infilling task

Take advantage of **bidirectional context** despite using **unidirectional models**

Please refer to the paper for more experiments and detailed analysis

Takeaways

Conceptual simplicity

Minimal change to standard LM training

Model-agnostic framework

Leverage massively pre-trained LMs

Thank [blank] for [blank]!



Thank you for watching!

Input

Output



Paper <https://arxiv.org/abs/2005.05339>

Code <https://github.com/chrisdonahue/ilm>

Demo <https://chrisdonahue.com/ilm>