

Learning to Reason and Memorize with Self-Questioning

Jack Lanchantin, Shubham Toshniwal, Jason Weston, Arthur Szlam, Sainbayar Sukhbaatar



Overview

- Large language models struggle with multi-step reasoning and state-tracking
- We propose a method to solve both of these problems by allowing the model to ask and answer questions
- The model can deviate from the input context at any time for self-questioning
- Allows the model to recall info and reason on the fly as it reads context, extending its memory and enabling multi-step reasoning

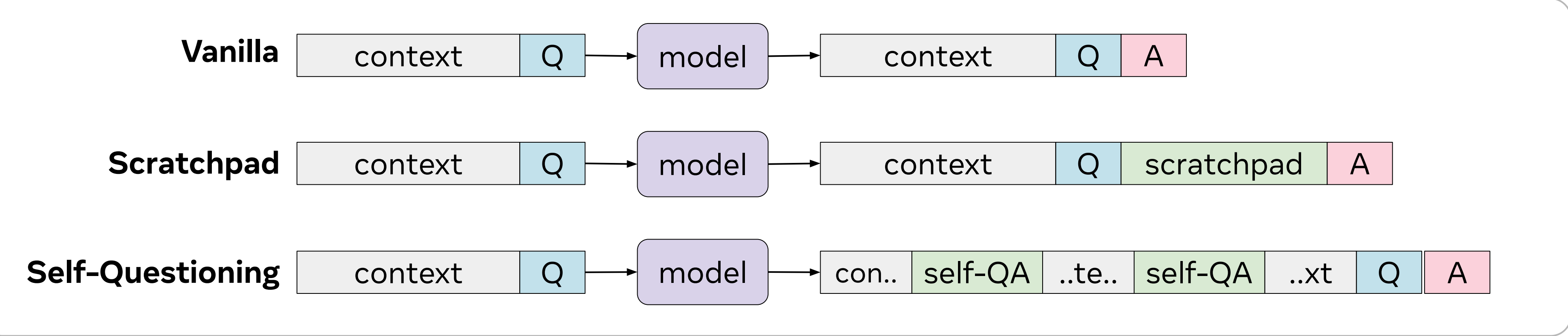
Method: Self-questioning

2 hop question:
Alice has the key. Alice is at the lake.
Q: where is the key? The key is at the lake.

3 hop question:
The key is in the bag. Alice has the bag. Alice is at the lake.
Q: where is the key? The key is at the bag.

3 hop question with self-questioning:
The key is in the bag. Alice has the bag. SQ: who has the key? Alice has the key. Alice is at the lake.
Q: where is the key? The key is at the lake.

Model



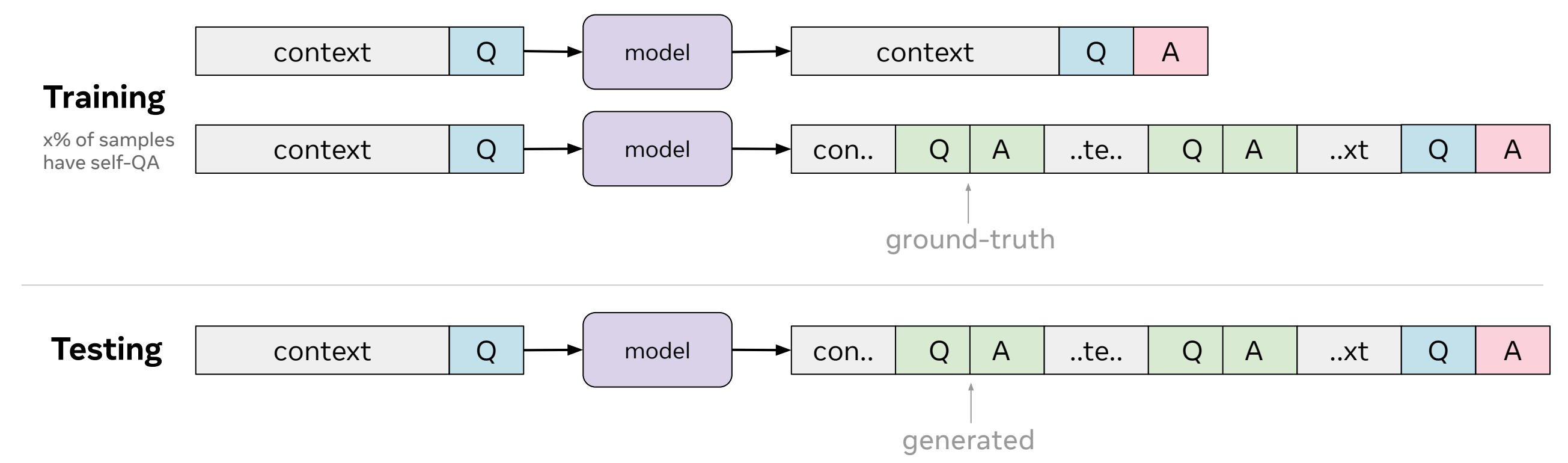
Data

Algorithmic Task
Original context:
e = 3;
e ++;
i = 4;
e --;
if i > e: e --;
g = 3;
Question: print e
Answer: e = 2;
Context with self-QA:
e = 3;
print e e = 3;
e ++;
print e e = 4;
i = 4;
print i i = 4;
e --;
print e e = 3;
if i > e: e --;
print e e = 2;
g = 3;
print g g = 3;
Question: print e
Answer: e = 2;

Toy-story Task
Original context:
The banana is inside the box.
Jessie has the bag.
The ball is inside the box.
The key is inside the suitcase.
Sid has the box.
Buzz has the suitcase.
Woody is at the station.
Question: who has the key?
Answer: Buzz has the key.
Context with self-QA:
The banana is inside the box.
Jessie has the bag.
The ball is inside the box.
The key is inside the suitcase.
Sid has the box.
SQ: Who has the banana?
Sid has the banana.
SQ: Who has the ball?
Sid has the ball.
Buzz has the suitcase
Woody is at the station.
Question: who has the key?
Answer: Buzz has the key.

Green tokens: self-QAs. Red tokens: final answer. "print" and "SQ:" are special tokens to generate a self-QA. If the model predicts the special token as the next token, it is allowed to generate a question and answer before returning to the original context.

Training and Testing



Results

Method	Train self-QA	Test self-QA	Algorithmic			Toy-Story	
			≤50	≤100	≤200*	3-hop*	4-hop*
GPT-2 fine-tuned	none	none	69.5	53.2	40.0	59.8	39.4
Scratchpad	100%	generated	100.0	61.0	30.0	98.6	95.4
Self-Questioning	100%	generated	100.0	100.0	98.8	99.7	98.6
	75%		99.5	99.6	98.3	98.9	98.5
	50%		98.8	98.2	95.0	98.9	96.8
	25%		95.8	93.5	85.7	97.1	95.1
Ablation	100%	100%	100.0	100.0	100.0	100.0	100.0
	100%	none	25.0	16.3	15.9	39.7	32.5

Test Accuracy. Train and test columns show what % of samples in the split have access to ground truth self-QAs, or if they are generated by the model itself. "*" indicates out-of-distribution test set.

Example output

Model	Context	Prediction
Vanilla (original context)	Mary is with Daniel. Frank is with Sandra. John has the book. Frank has the suitcase. Daniel is at the station. the banana is inside the basket. Bob has the apple. Bob has the bag. the ball is inside the box. the apple is inside the bag. Alice has the banana. Alice has the key. John is at the farm. Charlie is at the bridge. the book is inside the box. Alice is at the store. Bob is with Alice.	the ball is at the store.
Scratchpad (scratchpad context)	[SQ: Where is Mary? A: Mary is at the station. SQ: Who has the ball? A: Bob has the ball. SQ: Who has the basket? A: Alice has the basket. SQ: Where is the book? A: the book is at the farm. SQ: Where is the bag? A: the bag is at the ball. SQ: Where is the ball? A: the ball is at the bag. SQ: Who has the box? A: John has the box. SQ: Where is the box? A: the box is at the farm. SQ: Who has the key? A: Alice has the key. SQ: Where is the banana? A: the banana is at the store. SQ: Where is the basket? A: the basket is at the store. SQ: Where is the key? A: the key is at the store. SQ: Where is Bob? A: Bob is at the store. SQ: Where is the apple? A: the apple is at the store. SQ: Where is the suitcase? A: the suitcase is at the store. SQ: Where is the basket? A: the basket is at the store.]	the ball is at the ball.
Self-Questioning (self-QA context)	Mary is with Daniel. Frank is with Sandra. John has the book. Frank has the suitcase. Daniel is at the station. SQ: Where is Mary? Mary is at the station. the banana is inside the basket. Bob has the apple. Bob has the bag. the ball is inside the box. the apple is inside the bag. Alice has the banana. SQ: Who has the basket? Alice has the basket. Alice has the key. John is at the farm. SQ: Where is the book? the book is at the farm. Charlie is at the bridge. the book is inside the box. SQ: Who has the box? John has the box. SQ: Where is the box? the box is at the farm. SQ: Who has the ball? John has the ball. SQ: Where is the ball? the ball is at the farm. Alice is at the store. SQ: Where is the banana? the banana is at the store. SQ: Where is the basket? the basket is at the store. SQ: Where is the key? the key is at the store. Bob is with Alice. SQ: Where is Bob? Bob is at the store. SQ: Where is the apple? the apple is at the store. SQ: Where is the bag? the bag is at the store. SQ: Where is the key? the key is at the store.	the ball is at the farm.

tl;dr - a general method that allows language models to take internal notes in the form of self-QAs

Paper link:

