Stock QA via VAE with Retrieved Info

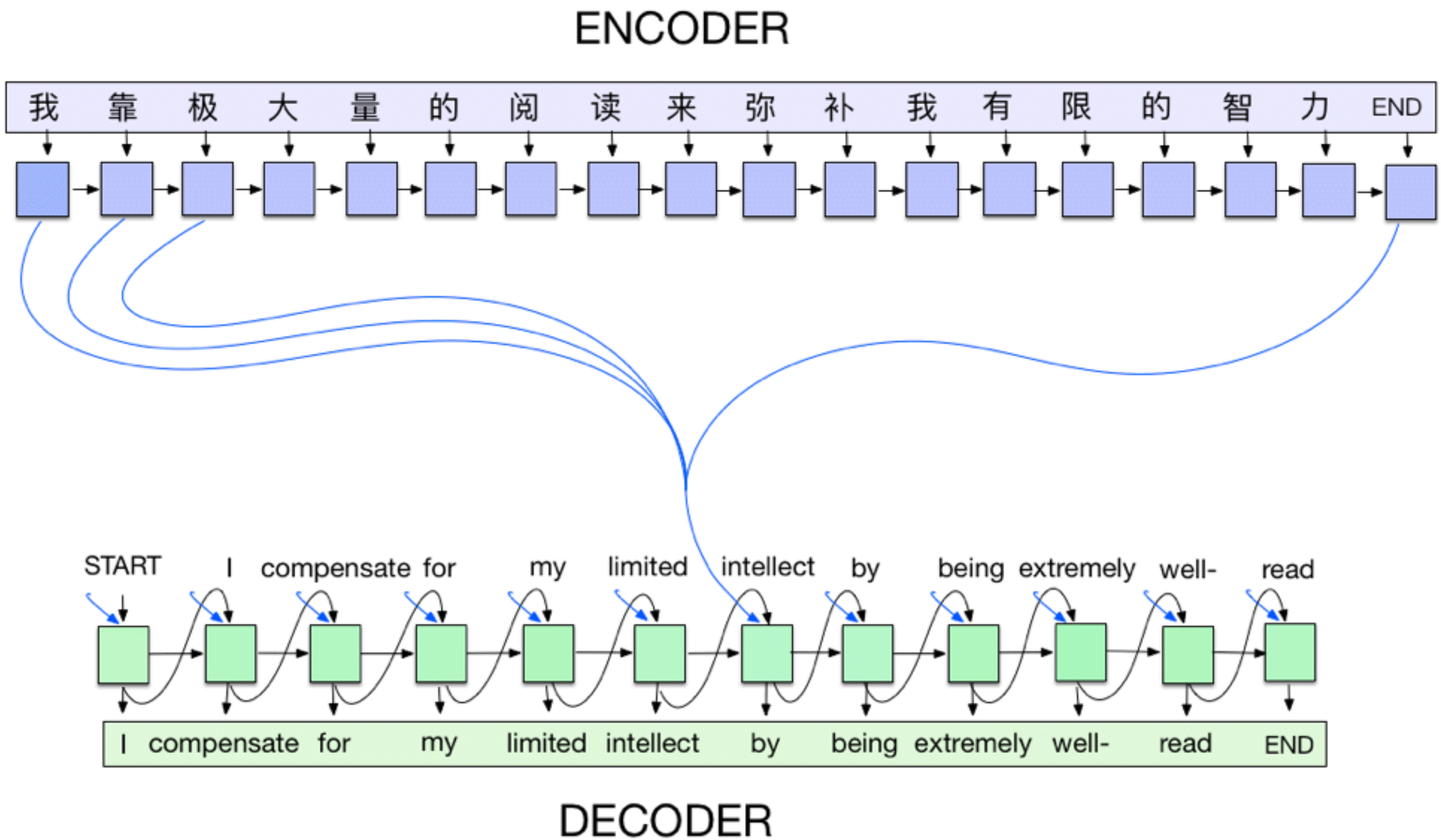
Yong Jiang

Stock QA Task

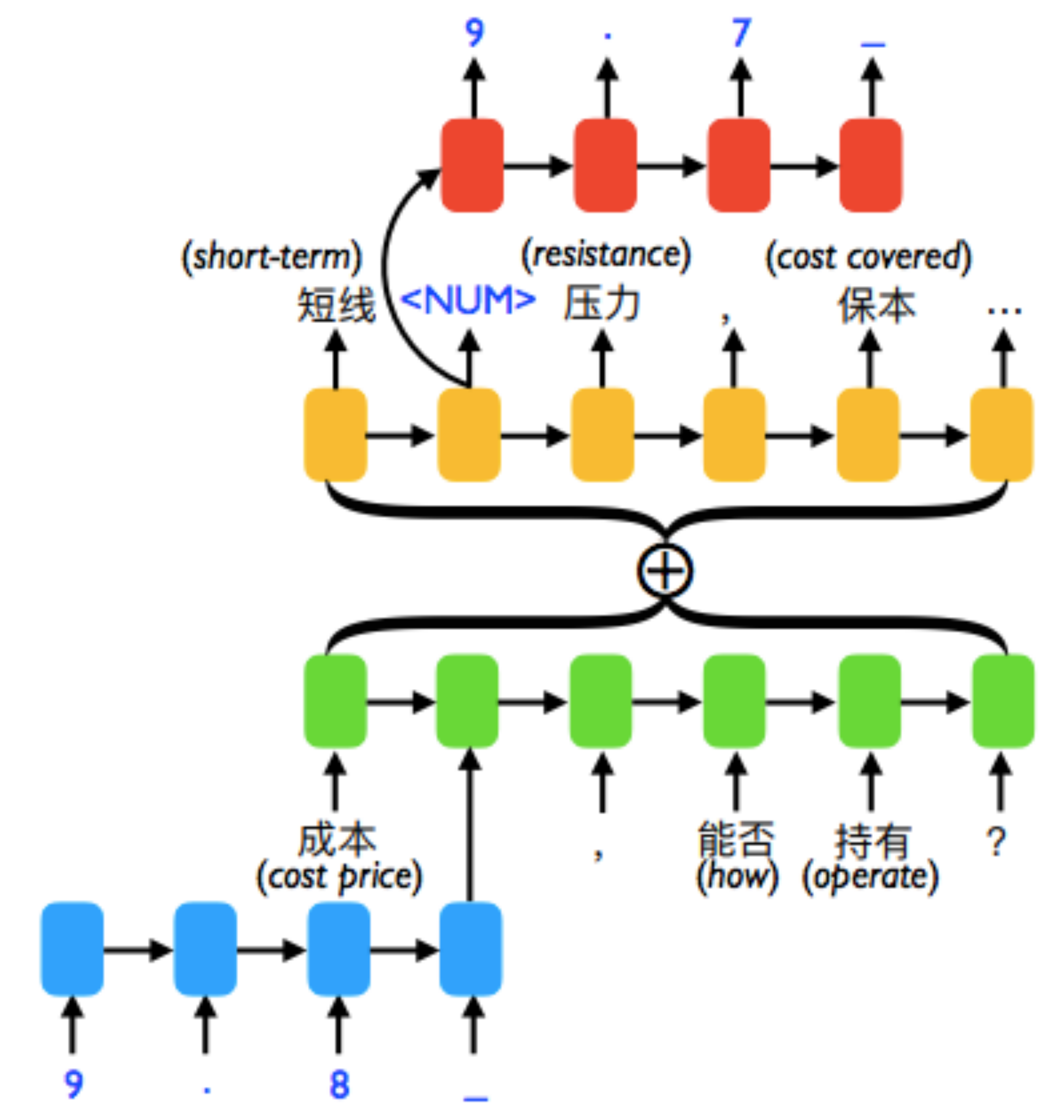
Q: I bought TSLA at $349, how to handle it?
A: TSLA is in a weakening trend. I suggest you to sell some shares to reduce your risk. Sell all if it could not rise above the 50-day MA $330.
Q: What is the support level of MSFT? I invested 30% of my money in it.
A: MSFT is trying to break the previous high point. Hold it if it can stay steady on $79.15.

Feature	Value	Feature	Value
Open	9.75	AvgVol5	53186.72
Close	9.15	AvgVol10	53186.72
High	9.93	Price change	-0.45
Low	9.02	Change rate	-0.05

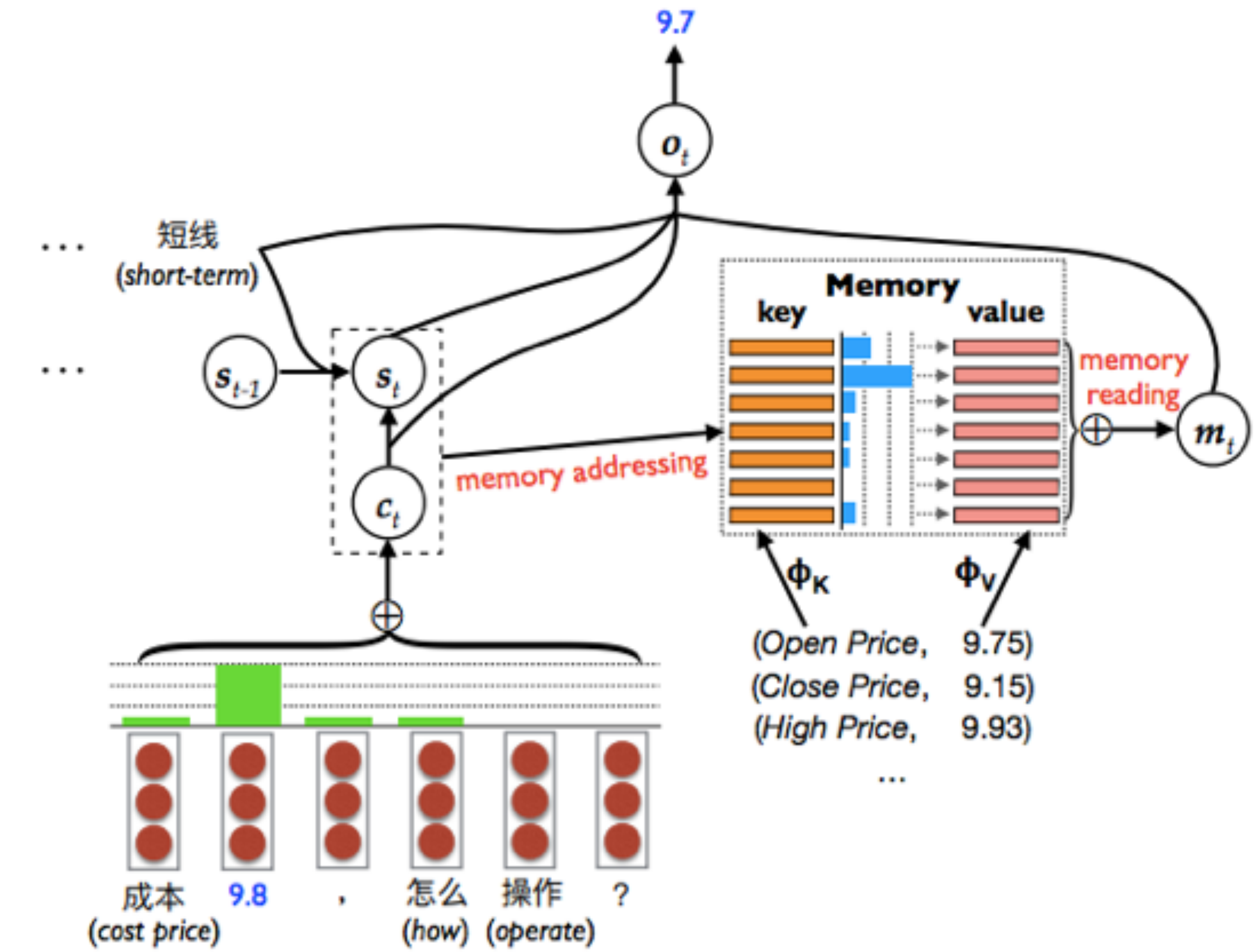
Encoder-Decoder Framework



Tu et al. 2018



hybrid word-character model

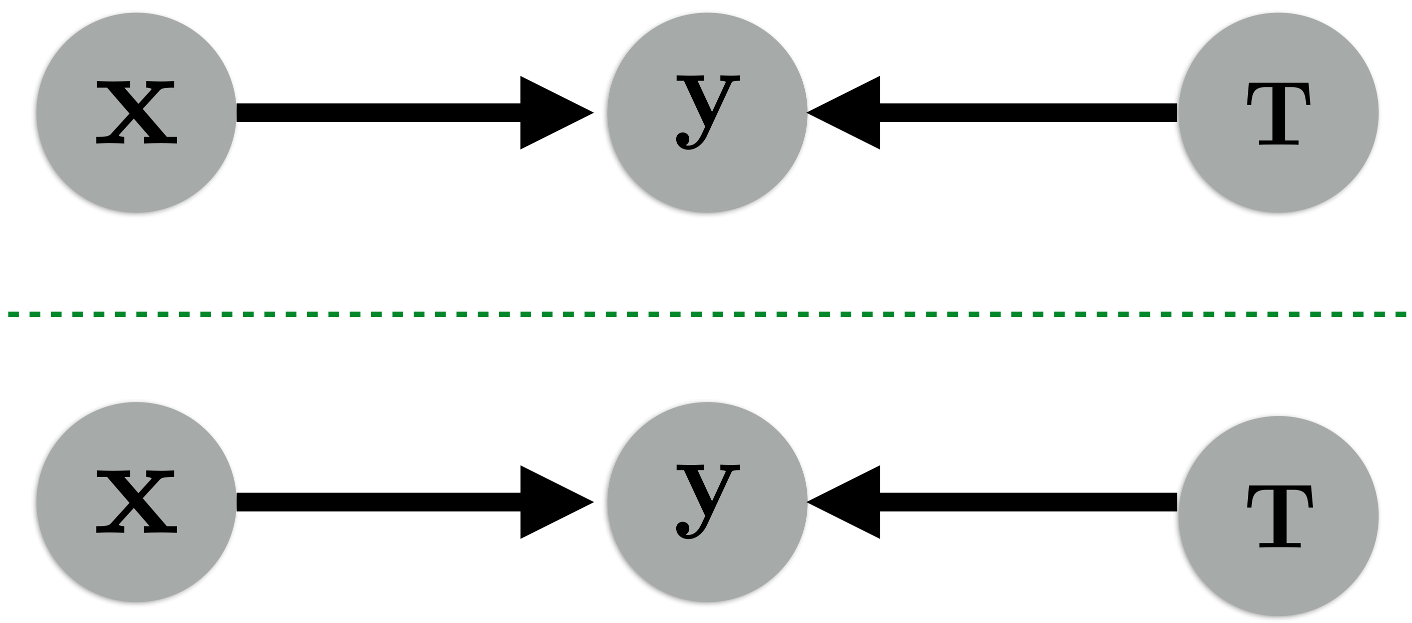


key-value memory net

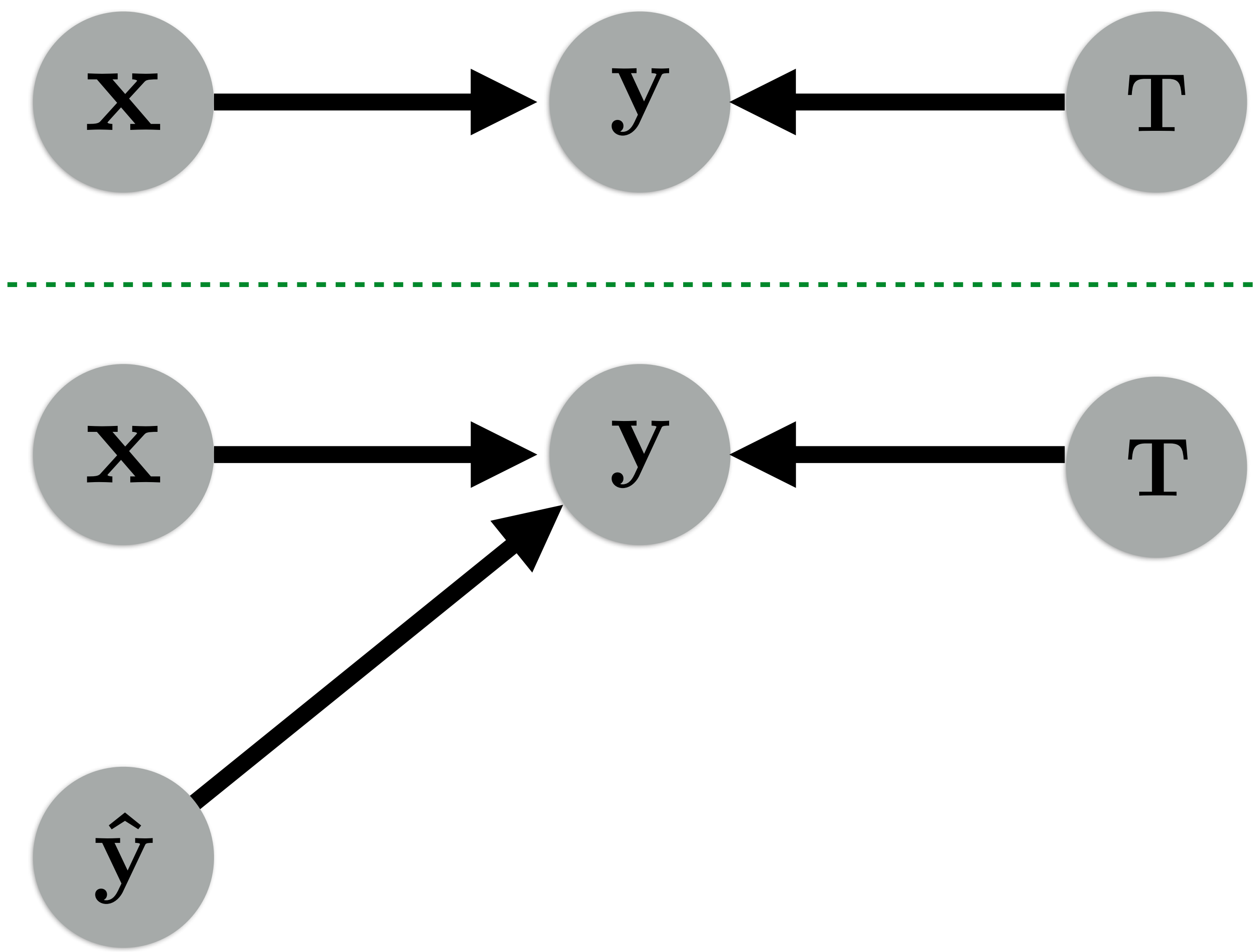
Weakness

- Generic responses
- Informativeness

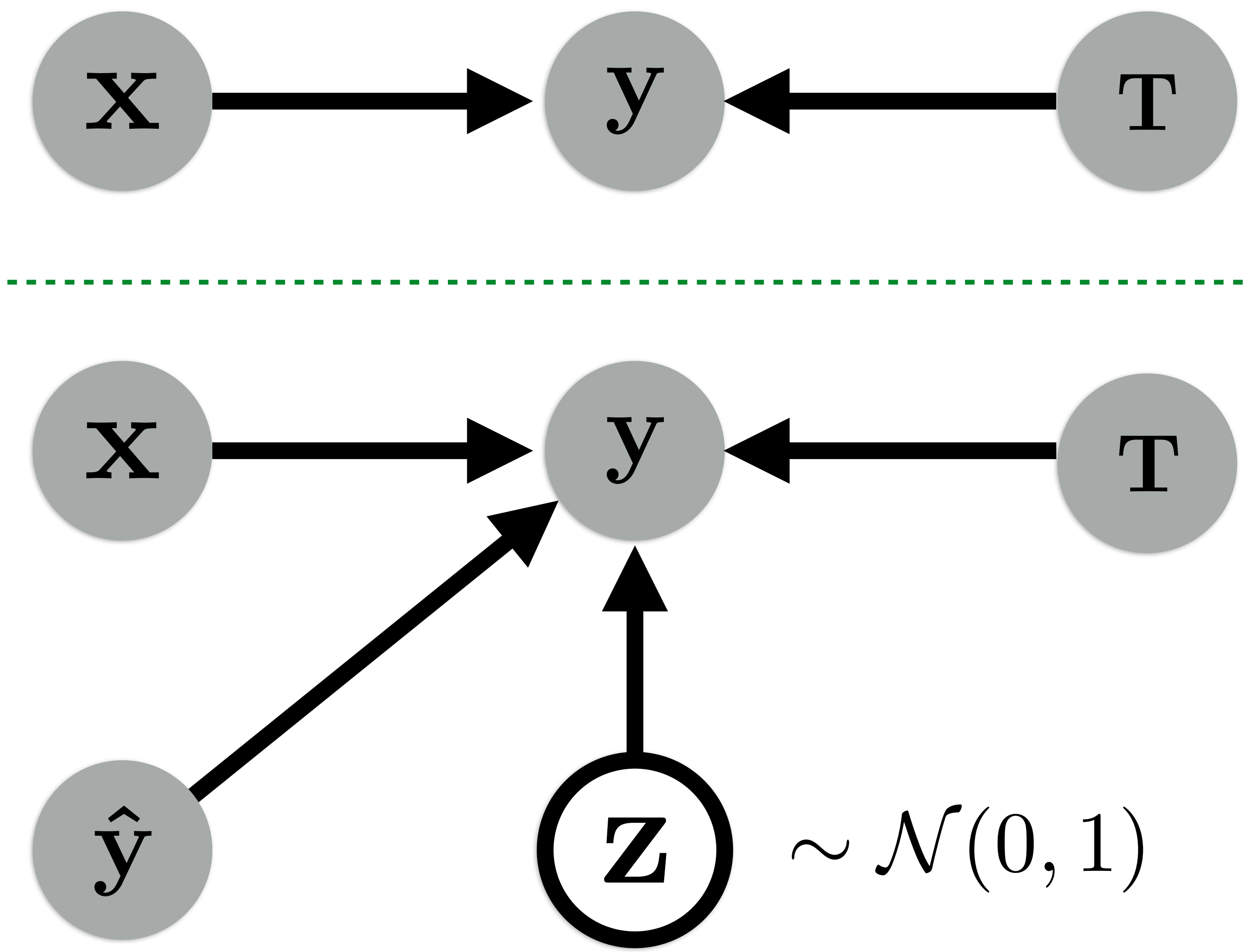
Generation with Retrievals



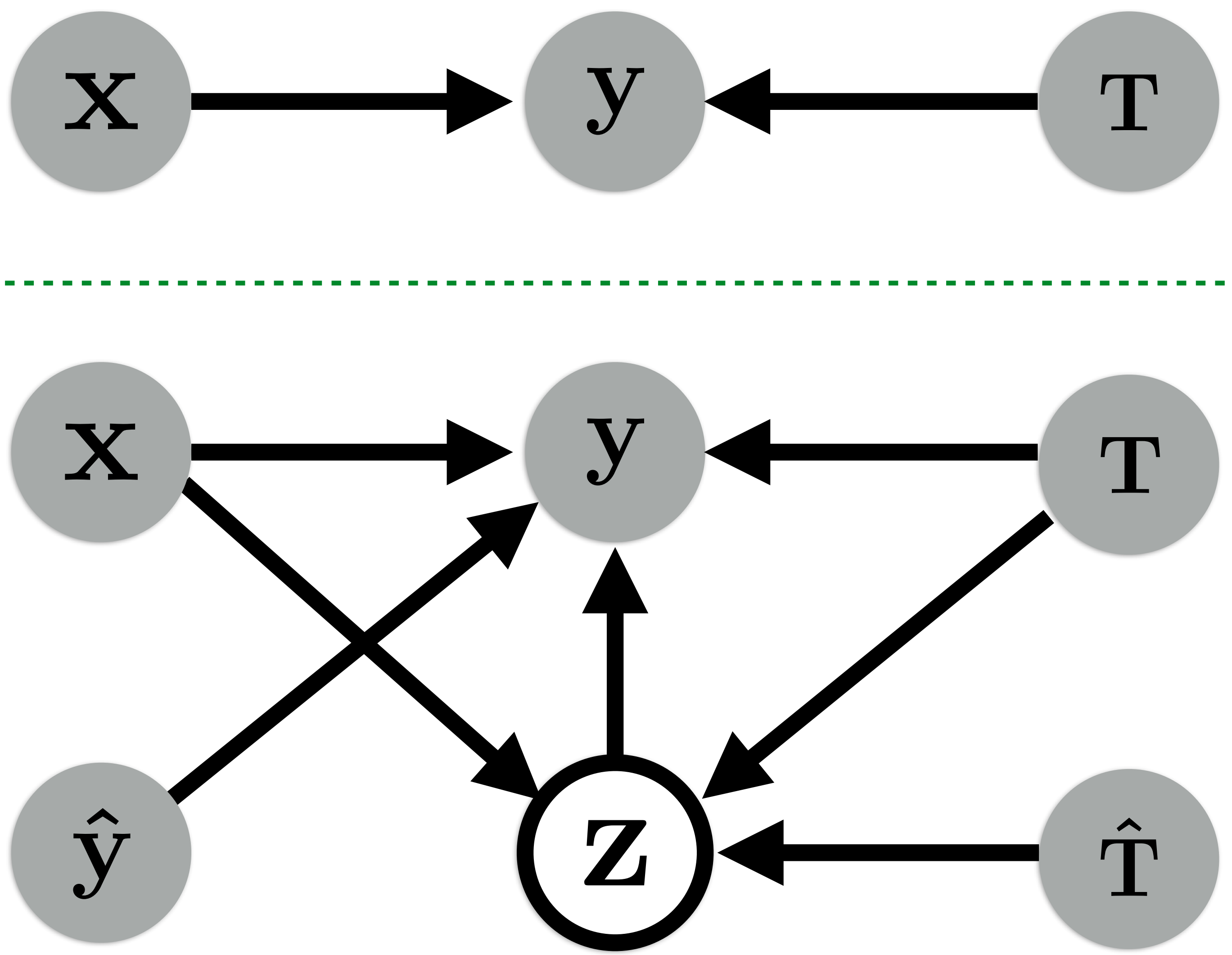
Generation with Retrievals



Generation with Retrievals

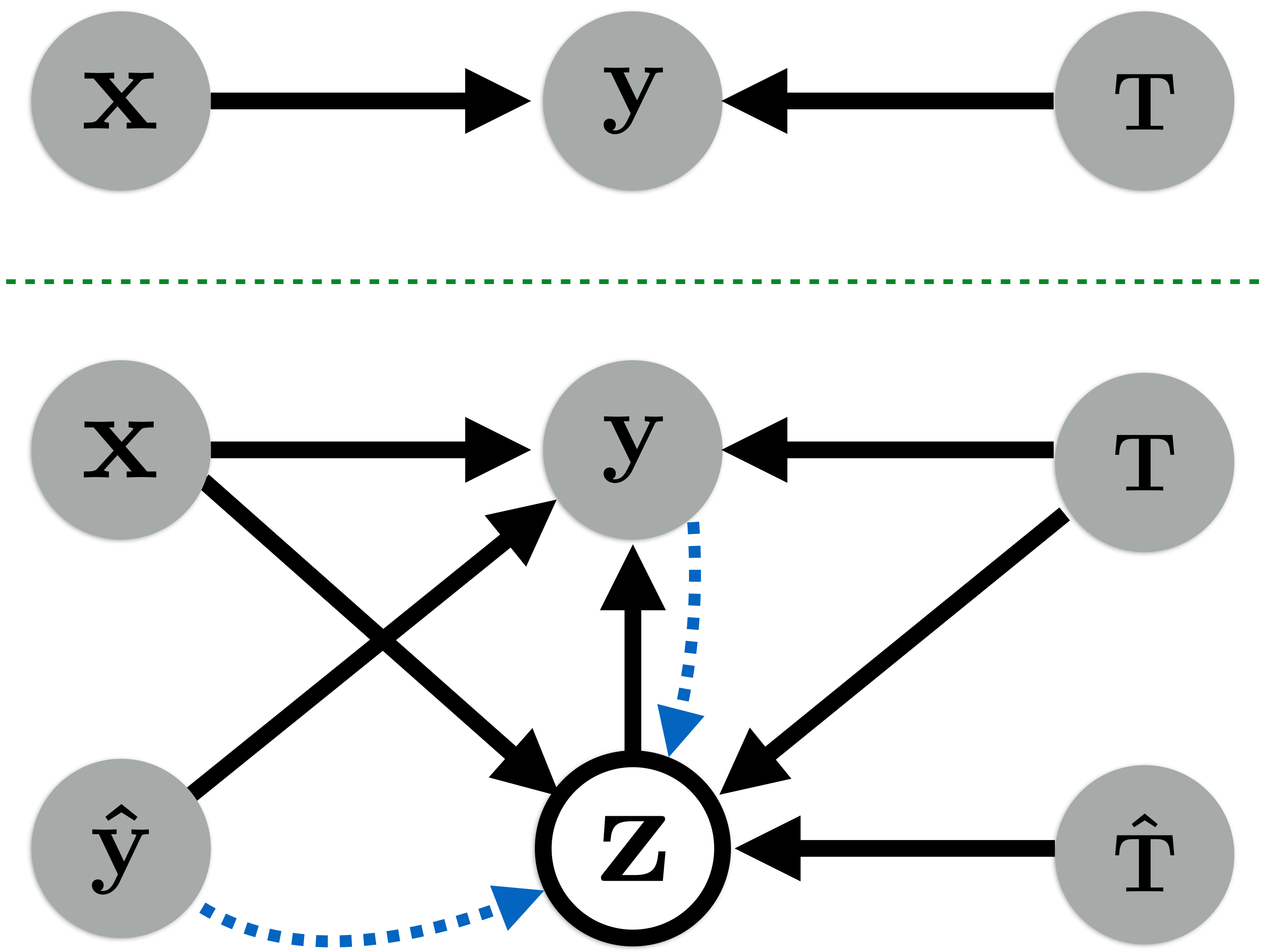


Generation with Retrievals

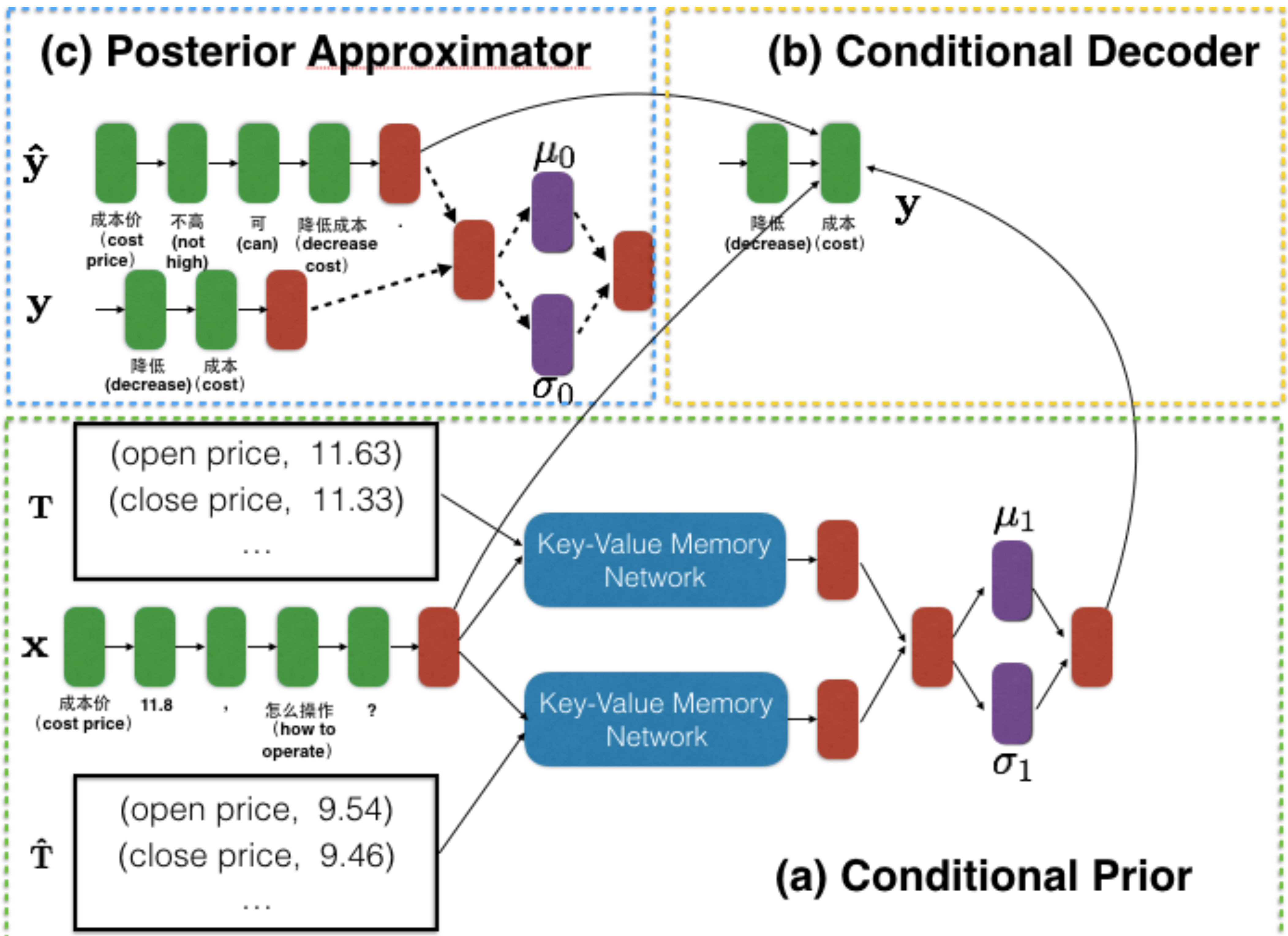


$$\sim \mathcal{N}(f(\mathbf{x}, \mathbf{T}, \hat{\mathbf{T}}), g(\mathbf{x}, \mathbf{T}, \hat{\mathbf{T}}))$$

Generation with Retrievals



Network Structure



Experiments

Model	Diversity at n-gram			
	1	2	3	4
Generative	0.12	0.27	0.37	0.42
Multi-Source	0.13	0.31	0.44	0.52
Edit Vector (Det)	0.13	0.32	0.46	0.54
Edit Vector (Diverse)	0.13	0.33	0.47	0.54
Retrieval	0.28	0.77	0.93	0.96

Table 4: Evaluation of diversity at different granularities, ranging from 1-gram to 4-gram. Higher number denotes higher diversity.

Model	Number of n-gram		
	1	2	3
Generative	345	639	797
Multi-Source	501	941	1178
Edit Vector (Det)	578	1186	1567
Edit Vector (Diverse)	596	1190	1577
Retrieval	1694	4182	4587

Table 5: Evaluation of informativeness at different granularities, ranging from 1-gram to 4-gram. Higher number denotes higher informativeness.

Current Problems

- Precision …
- Relation between retrieved answer and gold answer
- Repeated responses

Another Paper

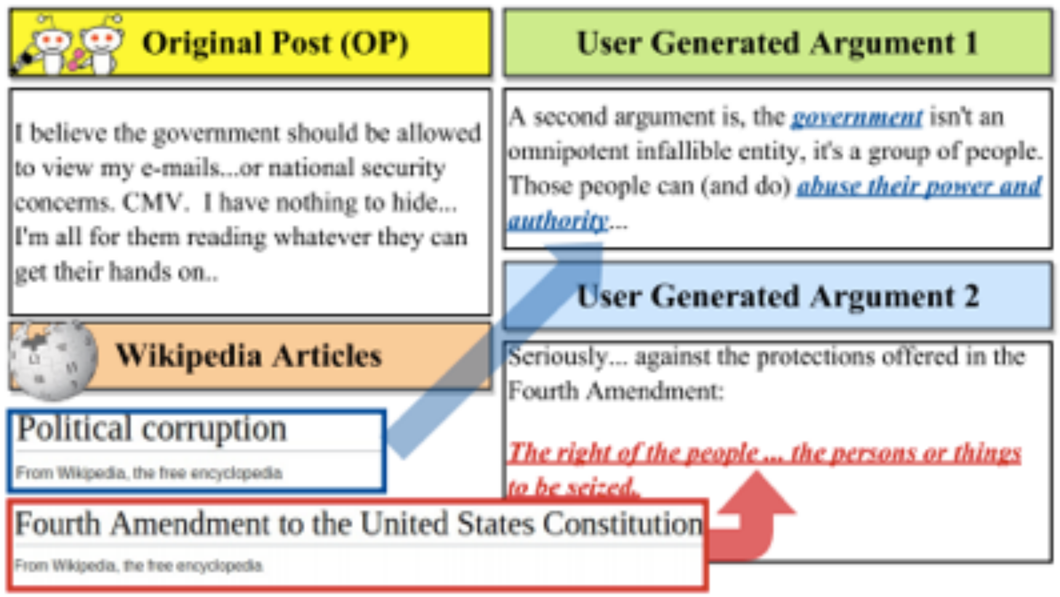
Neural Argument Generation Augmented with Externally Retrieved Evidence

Xinyu Hua and **Lu Wang**
College of Computer and Information Science
Northeastern University
Boston, MA 02115
hua.x@husky.neu.edu luwang@ccs.neu.edu

5 May 2018

Abstract

High quality arguments are essential elements for human reasoning and decision-making processes. However, effective argument construction is a challenging task for both human and machines. In this work, we study a novel task on *automatically generating arguments of a different stance for a given statement*. We propose an encoder-decoder style neural network-based argument generation model en-



Another Paper

Framework-1

- Pipeline: Evidence retrieval + Argument construction

Framework-1

- Pipeline: Evidence retrieval + Argument construction
- Two steps for retrieving
  - Extract NP/VP as one query for each sentence
  - Top five retrieved articles with highest TF-IDF similarity scores are kept
  - Articles -> Segmented paragraphs ->TOP 100 selected
  - Paragraphs -> Sentences -> TOP 10 selected

Framework-1

- Pipeline: Evidence retrieval + Argument construction
- Two steps for retrieving
  - Extract NP/VP as one query for each sentence
  - Top five retrieved articles with highest TF-IDF scores are kept
  - Articles -> Segmented paragraphs ->TOP 100 selected
  - Paragraphs -> Sentences -> TOP 10 selected
- Two decoders for generation
  - Keyphrase decoder
  - Argument decoder

Framework-2

- Different queries from training & testing
  - Training: queries are constructed from target argument
  - Testing: queries are constructed from input argument

Framework-2

- Different queries from training & testing
  - Training: queries are constructed from target argument
  - Testing: queries are constructed from input argument
- Keyphrase Construction
  - NP/VP constructed using CoreNLP
  - Keep NP/VP of length 2-10
  - If Keyphrases are overlap:
    - Longer is kept if it has more content word coverage
    - Shorter is kept otherwise

Network Structure

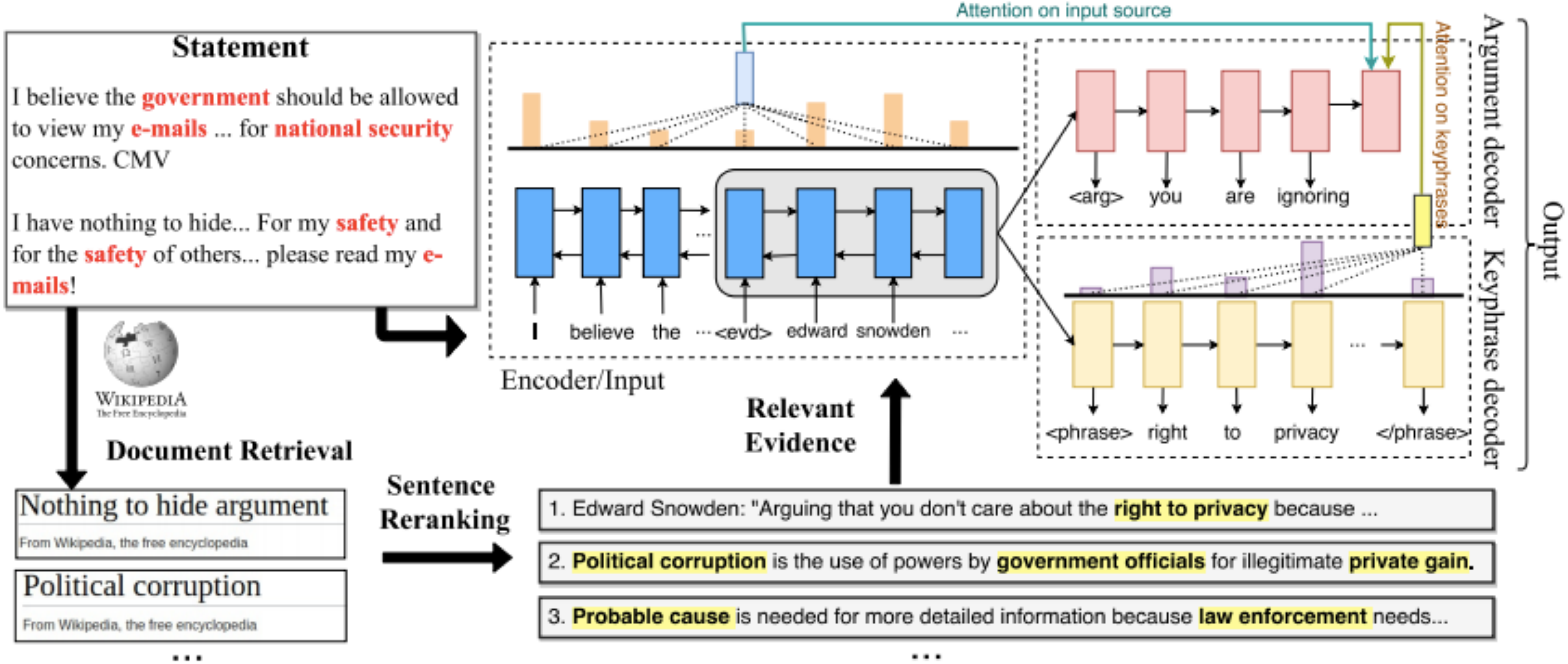


Figure 2: Overview of our system pipeline (best viewed in color). Given a statement, relevant articles are retrieved from Wikipedia with topic signatures from statement as queries (marked in red and boldface). A reranking module then outputs top sentences as evidence. The statement and the evidence (encoder states in gray panel) are concatenated and encoded as input for our argument generation model. During decoding, the keyphrase decoder first generates talking points as phrases, followed by the argument decoder which constructs the argument by attending both input and keyphrases.

Experiments

	w/ System Retrieval			*w/ Oracle Retrieval*		
	BLEU	**MTR**	**Len**	**BLEU**	**MTR**	**Len**
Baseline						
RETRIEVAL	15.32	**12.19**	151.2	10.24	**16.22**	132.7
Comparisons						
SEQ2SEQ	10.21	5.74	34.9	7.44	5.25	31.1
+ *encode evd*	18.03	7.32	67.0	13.79	10.06	68.1
+ *encode KP*	21.94	8.63	74.4	12.96	10.50	78.2
Our Models						
DEC-SHARED	21.22	8.91	69.1	15.78	11.52	68.2
+ *attend KP*	**24.71**	10.05	74.8	11.48	10.08	40.5
DEC-SEPARATE	24.24	10.63	88.6	17.48	13.15	86.9
+ *attend KP*	24.52	11.27	88.3	**17.80**	13.67	86.8

Table 3: Results on argument generation by BLEU and METEOR (MTR), with system retrieved evidence and oracle retrieval. The best performing model is highlighted in **bold** per metric. Our separate decoder models, with and without keyphrase attention, statistically significantly outperform all seq2seq-based models based on approximation randomization testing (Noreen, 1989), $p < 0.0001$.