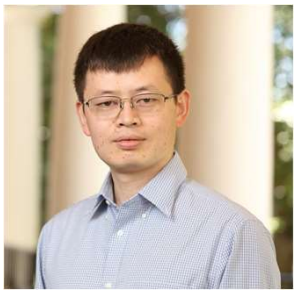
HUGGING FACE

The Amazing World of Neural Language Generation

EMNLP 2020

November 20, 2020



Yangfeng Ji
University Of Virginia

Antoine Bosselut
Stanford University

Thomas Wolf
Huggingface

Asli Celikyilmaz
Microsoft Research

Natural Language Generation

- Sub-field of natural language processing (NLP)
- Building software systems to produce:
  - ***coherent***, ***readable*** and **useful** written or spoken text.
- Produces explanations, summaries, narratives, etc.

Machine Translation

[User] We will discuss several issues today!

[System] Nous discuterons plusieurs questions aujourd'hui.

[System] 我们今天将讨论几个问题！

[System] Bugun cok sayıda sorunu tartısacağız.

Conversational Dialog Systems

[USER] Where is my next appointment and am I free for lunch?

[Agent] Your next meeting is at 10:30 at City Center. Did you want me to book a place for lunch in downtown ?

Understanding Visually Grounded Language

Image Source : bing.com

Image or Video Captioning

a man is showing how to make a drink

A car drives really fast

Olivastri, et.al., ICCVW, 2019

Multi Document Summarization

Summary

High Quality Content by WIKIPEDIA articles! Multi-document summarization is an automatic procedure aimed at extraction of information from multiple texts written about the same topic. Resulting summary report allows individual users, so as professional information consumers, to quickly familiarize themselves with information contained in a large cluster of documents. In such a way, multi-document summarization systems are complementing the news aggregators performing the next step down the road of coping with information overload.

Image Source: Alphascript Publishing

Data-to-Text Generation

TEAM	WIN	LOSS	PTS	FG_PCT	RB	AST	...
Pacers	4	6	99	42	40	17	...
Celtics	5	4	105	44	47	22	...

PLAYER	H/V	AST	RB	PTS	FG	CITY	...
Jeff Teague	H	4	3	20	4	Indiana	...
Miles Turner	H	1	8	17	6	Indiana	...
Isaiah Thomas	V	5	0	23	4	Boston	...
Kelly Olynyk	V	4	6	16	6	Boston	...
Amir Johnson	V	3	9	14	4	Boston	...
...	...	...	...	...	...	...	

PTS: points, FT_PCT: free throw percentage, RB: rebounds, AST: assists, H/V: home or visiting, FG: field goals, CITY: player team city.

The **Boston Celtics** defeated the host **Indiana Pacers 105-99** at Bankers Life Fieldhouse on Saturday. In a battle between two injury-riddled teams, the Celtics were able to prevail with a much needed road victory. The key was shooting and defense, as the **Celtics** outshot the **Pacers** from the field, from three-point range and from the free-throw line. Boston also held Indiana to **42 percent** from the field and **22 percent** from long distance. The Celtics also won the rebounding and assisting differentials, while tying the Pacers in turnovers. There were 10 ties and 10 lead changes, as this game went down to the final seconds. Boston (**5–4**) has had to deal with a gluttony of injuries, but they had the fortunate task of playing a team just as injured here. **Isaiah** Thomas led the team in scoring, totaling **23 points and five assists on 4–of–13** shooting. He got most of those points by going 14–of–15 from the free-throw line. **Kelly Olynyk** got a rare start and finished second on the team with his **16 points, six rebounds and four assists**.

Puduppully, et.al.,2019

Meeting Summarization

C: Looking at what we've got, we we want an LCD display with a spinning wheel.
B: You have to have some push-buttons, don't you?
C: Just spinning and not scrolling, I would say.
B: I think the spinning wheel is definitely very now.
A: but since LCDs seems to be uh a definite yes,
C: We're having push-buttons on the outside
C: and then on the inside an LCD with spinning wheel,

Decision Abstract (Summary):
The remote will have push buttons outside, and an LCD and spinning wheel inside.

A: and um I'm not sure about the buttons being in the shape of fruit though.
D: Maybe make it like fruity colours or something.
C: The power button could be like a big apple or something.
D: Um like I'm just thinking bright colours.

Problem Abstract (Summary):
How to incorporate a fruit and vegetable theme into the remote.

Wang and Cardie, ACL, 2013

Conversational Dialog Summarization

[Nurse] Hi Mr.#name#, you were discharged on #date#. There are some questions i'd like to check with you.
[Patient] Ok, Ok.
[Nurse] Well, have you been experiencing swelling recently?
[Patient] Swelling? It comes and go, comes and go.
[Nurse] Comes and go ... I see .. **#repetition#**
[Nurse] ... **#pause#**... When did it start?
[Patient] Let me see, started from three weeks ago.
... ...
[Nurse] Are you experiencing any headache right now as we speak?
[Patient] Umm ... **#back-channel#**
[Nurse] Let me check, the last time you told me is sometimes at night.
[Patient] Oh, right, only a bit.
... ...
[Nurse] Still feel some chest pain or chest discomfort?
[Patient] Yes, my head is... **#false-start#** no, the pain is much better. Still feel headache though ... **#topic-drift#**
... ...
[Nurse] Any giddiness or palpitation?
[Patient] Palpitation? Do not have-- #**interruption**#
[Nurse] Well ... Do you-- #**interruption**#
[Patient] and no giddiness, no, nothing.
... ...
[Nurse] Ok, you need to check your heartrate everyday.
[Nurse] Do you know how to use the device?
[Patient] Yes, yes, no problem.
... ...

Swelling: started from three weeks ago, comes and go.
Headache: sometimes, at night, only a bit.
Chest pain: much better.
Dizziness: none.

Liu, et.al., ASRU 2019

Other Text Generation Tasks

Question Generation

Paraphrase Generation

Poetry Generation

Long Question Answering

Visual Dialog Systems

Document/Article Generation

Program Synthesis

Search Snippet Generation

Why Automatic Text Generation?

Efficiency

Why Automatic Text Generation?

Efficiency

Education

Why Automatic Text Generation?

Efficiency

Education

Productivity

History of Natural Langauge Generation Systems

Template based systems:
Uses rules and templates

1965

History of Natural Langauge Generation Systems

Template based systems:
Uses rules and templates

1965

1978

Rule-based + Data Driven pipelines :
Document planning + microplanning + realization

History of Natural Langauge Generation Systems

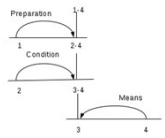
Template based systems:
Uses rules and templates

1965

1978

Rule-based + Data Driven pipelines :
Document planning + microplanning + realization

Modeling Discourse Structures :
Relation learning,
Rhetorical Structure Theory



1985

History of Natural Langauge Generation Systems

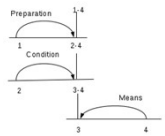
Template based systems:
Uses rules and templates

1965

1978

Rule-based + Data Driven pipelines :
Document planning + microplanning + realization

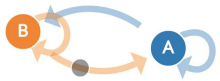
Modeling Discourse Structures :
Relation learning, Rhetorical Structure Theory



1985

1990

Statistical Methods (Markov Chains):
Sentence compression, reordering, lexical paraphrasing, syntactic transformation



History of Natural Langauge Generation Systems

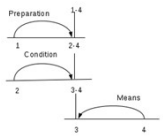
Template based systems:
Uses rules and templates

1965

1978

Rule-based + Data Driven pipelines :
Document planning + microplanning + realization

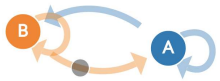
Modeling Discourse Structures :
Relation learning, Rhetorical Structure Theory



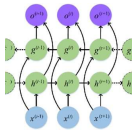
1985

1990

Statistical Methods (Markov Chains):
Sentence compression, reordering, lexical paraphrasing, syntactic transformation



RNNs, LSTMs, GRUs:
Autoregressive DNNs + Recurrent units, backpropagation



2013

History of Natural Langauge Generation Systems

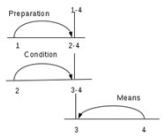
Template based systems:
Uses rules and templates

1965

1978

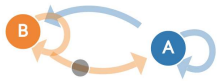
Rule-based + Data Driven pipelines :
Document planning + microplanning + realization

Modeling Discourse Structures :
Relation learning, Rhetorical Structure Theory



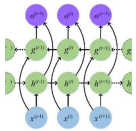
1985

Statistical Methods (Markov Chains):
Sentence compression, reordering, lexical paraphrasing, syntactic transformation



1990

RNNs, LSTMs, GRUs:
Autoregressive DNNs + Recurrent units, backpropagation



2013

2017

Transformers :
GPT (1/2/3), GROVER. TransformerXL, DialoGPT

History of Natural Langauge Generation Systems

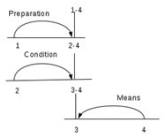
Template based systems:
Uses rules and templates

1965

1978

Rule-based + Data Driven pipelines :
Document planning + microplanning + realization

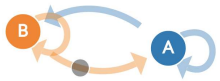
Modeling Discourse Structures :
Relation learning, Rhetorical Structure Theory



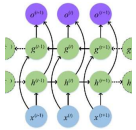
1985

1990

Statistical Methods (Markov Chains):
Sentence compression, reordering, lexical paraphrasing, syntactic transformation



RNNs, LSTMs, GRUs:
Autoregressive DNNs + Recurrent units, backpropagation



2013

2017

Transformers :
GPT (1/2/3), GROVER. TransformerXL, DialoGPT

Language Modeling

Conditional Language Modelling

$$P(w_1, \cdots, w_n) = \prod_{i=1}^{n} P(w_i | w_{i-1}, \cdots, w_1)$$

Examples from D. Jurafsky slides

Conditional Language Modelling

$$P(w_1, \cdots, w_n) = \prod_{i=1}^{n} P(w_i | w_{i-1}, \cdots, w_1)$$

- Speech Recognition
  - P(I saw a man) >> P(eyes awe of an)

Examples from D. Jurafsky slides

Conditional Language Modelling

$$P(w_1, \cdots, w_n) = \prod_{i=1}^{n} P(w_i | w_{i-1}, \cdots, w_1)$$

- Speech Recognition
  - P(I saw a man) >> P(eyes awe of an)
- Spell Correction
  - P(about fifteen minutes from) > P(about fifteen minuets from)

Examples from D. Jurafsky slides

Conditional Language Modelling

$$P(w_1, \cdots, w_n) = \prod_{i=1}^{n} P(w_i | w_{i-1}, \cdots, w_1)$$

- Speech Recognition
  - P(I saw a man) >> P(eyes awe of an)
- Spell Correction
  - P(about fifteen minutes from) > P(about fifteen minuets from)
- Machine translation, Question Answering, Paraphrasing, Image captioning, Summarization, others…

Examples from D. Jurafsky slides

Long Text Generation – Image Story Telling

Two men are riding bicycles down a street. One man is behind the other riding on the side of a paved road. Both bicycles have black baskets on the front attached to the handlebars. Both men are holding umbrellas. The first man is wearing blue pants and a white and black striped shirt. His umbrella is blue. The second man is wearing black pants and a light purple shirt. His umbrella is aqua blue colored. Behind them, along the road are shops. One of the shops is closed with a silver metal fence covering the entrance. The other shop is a clothing store and there are four mannequins with clothing outside the door. There is also a pot-ted plant with red flowers outside one of the shops.

Krause, et.al., CVPR, 2017

Long Text Generation – Image Story Telling

Two men are riding bicycles down a street. One man is behind the other riding on the side of a paved road. Both bicycles have black baskets on the front attached to the handlebars. Both men are holding umbrellas. The first man is wearing blue pants and a white and black striped shirt. His umbrella is blue. The second man is wearing black pants and a light purple shirt. His umbrella is aqua blue colored. Behind them, along the road are shops. One of the shops is closed with a silver metal fence covering the entrance. The other shop is a clothing store and there are four mannequins with clothing outside the door. There is also a pot-ted plant with red flowers outside one of the shops.

Krause, et.al., CVPR, 2017

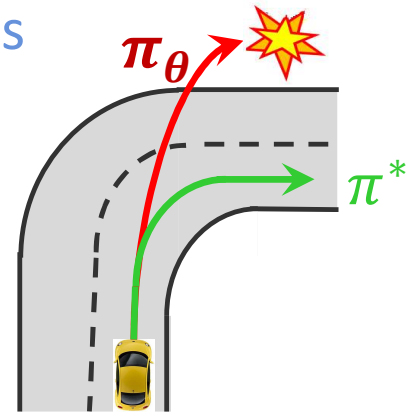
Long Text Generation with RNN-LM

two men are **riding a bicycle** on the road. it is a sunny day. a man is **riding a blue bicycle** on the street. the men with blue umbrella is **riding a bicycle** on a rainy day. the woman black dress is standing on the road.

Neural Text Generation Issues – Model Selection

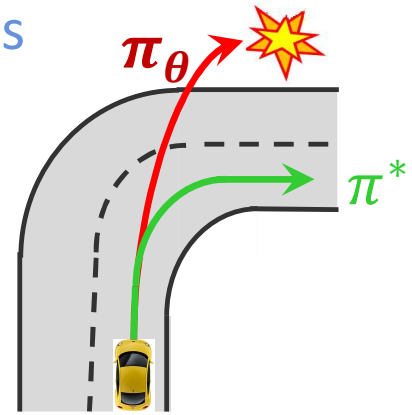
Neural Text Generation Issues – Model Selection

Exposure Bias



Neural Text Generation Issues – Model Selection

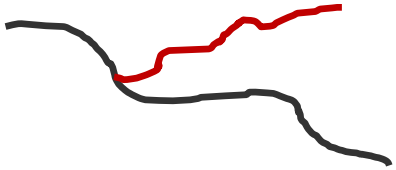
Exposure Bias



Compounding Errors / Label Bias

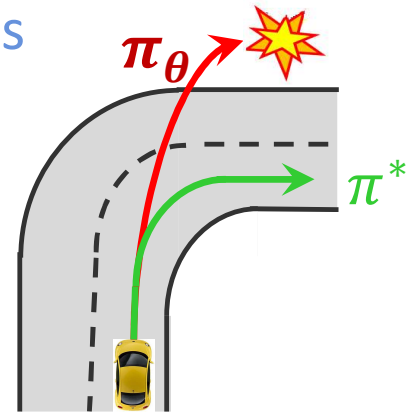
Gold: The cat purrs

Pred: The dog barks



Neural Text Generation Issues – Model Selection

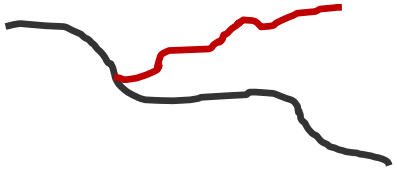
Exposure Bias



Compounding Errors / Label Bias

Gold: The cat purrs

Pred: The dog barks



Coherence and Narrative Flow

"do you like animals?"

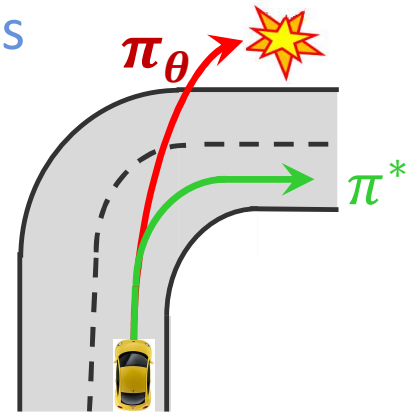
yes, I have three cats.

how many cats do you have ?

I don't have cats.

Neural Text Generation Issues – Model Selection

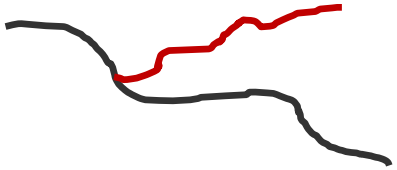
Exposure Bias



Compounding Errors / Label Bias

Gold: The cat purrs

Pred: The dog barks



Coherence and Narrative Flow

"do you like animals?"

yes, I have three cats.

how many cats do you have ?

I don't have cats.

Surrogate objective function

$$P^*(x|c) \rightarrow P_\theta(x|c)$$

Perplexity

vs

BLEU/ROUGE/Neural Rewards

Weaknesses of MEGA Language Models for GENERATION!

Inconsistent output

Crippled by length

coreference issues

Longer string that are repeated many times in the dataset

Unnecessarily repeating entities

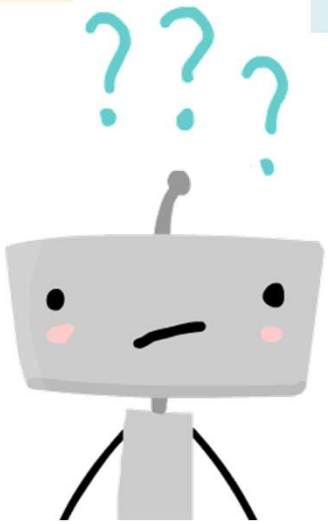
Maintaining **coherence** between **paragraphs**

No **real** understanding, **commonsense**, **factual correctness**"

How to learn discourse?

Lack of implicit "**planning**"

sub-optimal evaluation metrics "perplexity"



MLE: cannot capture high-level semantics

Biased pre-trained models

Domain transfer is hard

Single path generative flow

Softmax Bottleneck issues!

Surrogate-loss functions

Auto-regressive!

exposure-bias

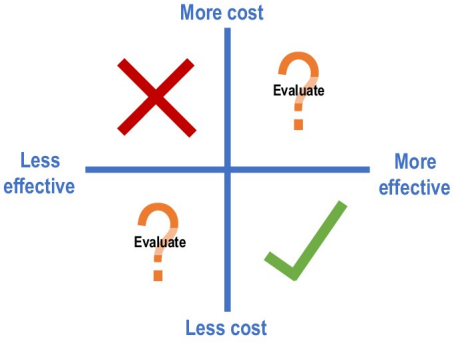
Degenerative sampling methods

word-by-word generation: can't see **global context**!

Research on Evaluation: Now more than ever!

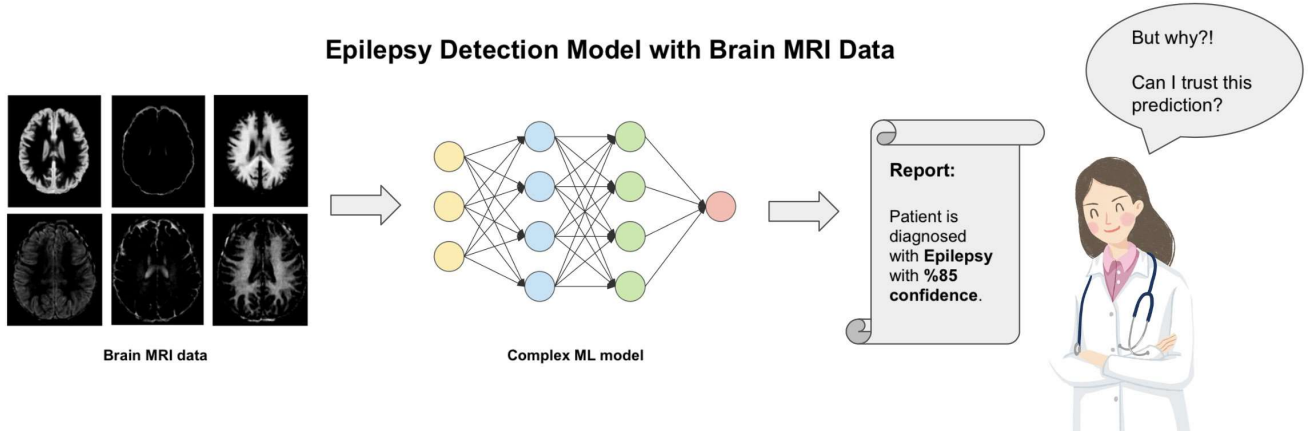
Detecting machine generated text

Standards in Text Generation Evaluations



Effective Evaluation

Evaluating Ethical Issues



Making Evaluation Explainable

Tutorial Schedule

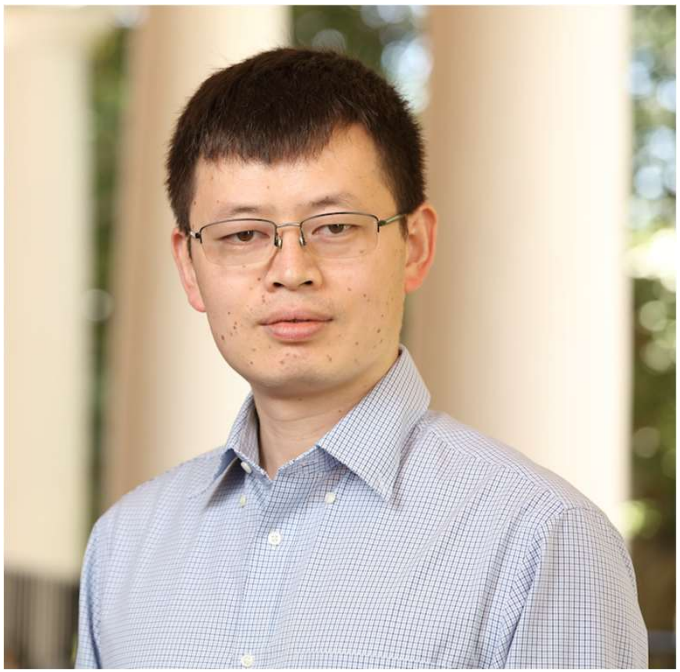
Approximate Time	Segment Title	Speaker	Description
15 mins	Introduction	Asli Celikyilmaz	Why is it important today to discuss neural text generation?
20 mins	Modeling: NN Architecture	Yangfeng Ji	Journey of NNs architectures that are used for text generation up-to-date
20 mins	Modeling: Generation With Rich context	Yangfeng Ji	How to efficiently use context in langauge generation?
20 mins	Training Text Generation Models	Antoine Bosselut	What are the best practices in training neural text generation systems today?
20 mins	Decoding Algorithms	Antoine Bosselut	Amazing world of neural decoding algorithms and beyond!
35 mins	Evaluation and Benchmarks	Asli Celikyilmaz	How are neural text generation models evaluated today? Future directions!
40 mins	Building NLG Systems	Thomas Wolf	Adventure into building an NLG system using state-of-the art tools and libraries.

Introduction and Evaluation

"Let's discuss the Evaluation of Text Generation Systems, which is more important today than ever!"

Time:

Modeling



Training and Decoding

Building of NLG Systems