

A NEURAL CONVERSATIONAL MODEL

Oriol Vinyals, Quoc V. Le

INTRODUCTION

- Objective: Conversational modeling i.e. developing systems that converse with humans, closely imitating natural human interactions.
- Limitations of Previous Approaches:
 - Domain-Specific: Prior models were often restricted to specific domains, such as booking airline tickets.
 - Rule-Heavy: Relied heavily on handcrafted rules.
- Goal:
 - Reduce Complexity: Minimize the need for feature engineering and domain specificity.
 - Enhance Performance: Aim to match or surpass current state-of-the-art results with a more generalized framework.

DATASETS

- IT Helpdesk Troubleshooting Dataset -
 - Typically 400 words per interaction.
 - Turn taking clearly indicated.
 - Dataset Size:
 - Training Set: 30 million tokens
 - Validation Set: 3 million tokens
- OpenSubtitles Dataset -
 - Assumption - Consecutive sentences are uttered by different characters.
 - Turn taking not clearly indicated.
 - Dataset Size:
 - Training Set: 62 million sentences (923 million tokens).
 - Validation Set: 26 million sentences (395 million tokens).

MODELING APPROACH

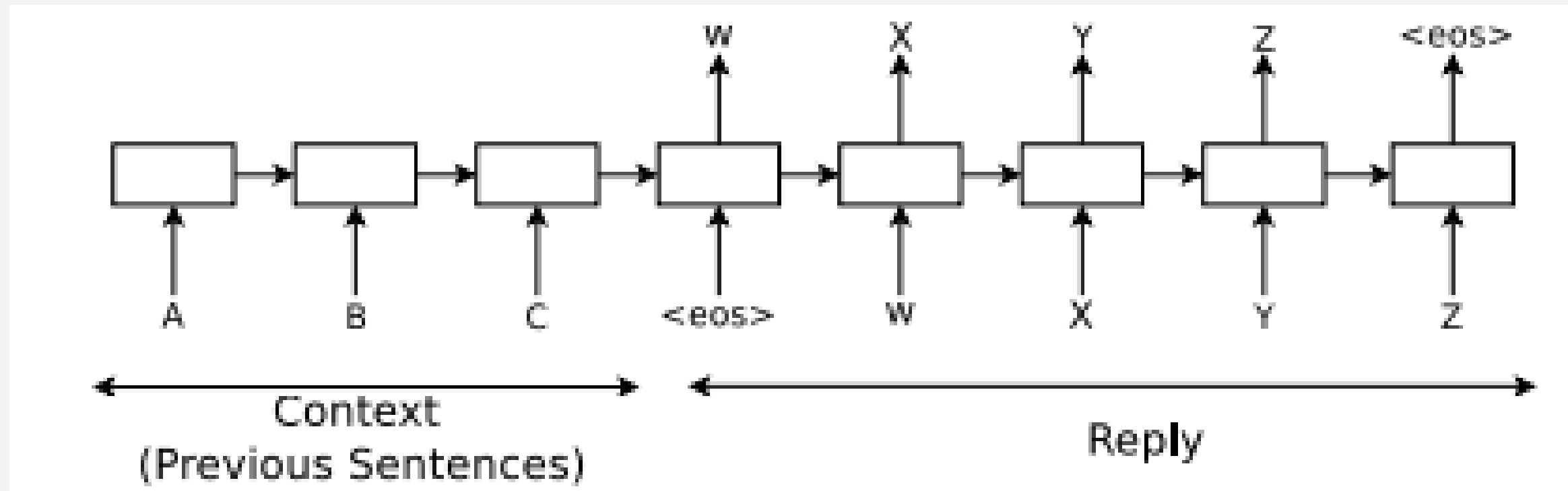


Figure 1. Using the seq2seq framework (Sutskever et al., 2014) for modeling conversations.

- Transforms input sequences into output sequences via an encoder-decoder architecture.
- It converses by predicting the next sentence based on the previous sentence or sentences in the conversation.

EXPERIMENTS

- IT Helpdesk Troubleshooting Experiment -
 - Model Configuration:
 - Vocabulary size = 20K
 - Single layer LSTM with 1024 memory cells.
 - Stochastic gradient descent with gradient clipping.
 - Perplexity = 8.
- OpenSubtitles experiment -
 - Model Configuration:
 - Vocabulary size = 100K
 - Two-layered LSTM, each layer containing 4096 memory cells.
 - AdaGrad with gradient clipping.
 - Perplexity = 17.

SAMPLE CONVERSATION

- Conversation 1: Browser Issues

Describe your problem: browser keeps crashing

Machine: *hi*

Human: *hi*

Machine: *could you please let me know what is the operating system you are using ?*

Human: *linux*

Machine: *ok this is the issue with the system performance , nasser , tools .*

Human: *what is that ?*

Machine: *is it a windows machine ?*

Machine: *we have linux on your machine .*

Machine: *is it windows xp or windows 7 ?*

Human: *i am using linux*

Machine: *let me take a remote of your machine and check if issue has been pushed properly .*

Human: *ok*

Machine: *please open terminal and run the below commands in one computer :*

Machine: *sudo apt - get update*

Machine: *sudo apt - get upgrade*

Machine: *sudo <COMMAND> - updater*

- Conversation 2: Password Issues

Describe your problem: i forgot my password

Machine: *hi , this is <NAME> from helpdesk connect , what can i help you with today ?*

Human: *hi , i forgot my password*

Machine: *ok great*

Human: *can you tell me how i can recover it ?*

Machine: *i ' ll need to verify who it is at that machine*

Machine: *can we do a hangout*

Machine: *?*

Human: *yes*

Machine: *<anonymized hangout instructions and verifications>*

Human: *cool , i am good now*

Machine: *how ' s this looking good ?*

Machine: *i ' ll go ahead and close this chat .*

Machine: *take care . :)*

Human: *bye*

Machine: *bye .*

SAMPLE CONVERSATION

- Conversation 1: General Knowledge Q&A

Human: *who is skywalker ?*

Machine: *he is a hero .*

Human: *who is bill clinton ?*

Machine: *he 's a billionaire .*

Human: *is sky blue or black ?*

Machine: *blue .*

Human: *does a cat have a tail ?*

Machine: *yes .*

Human: *does a cat have a wing ?*

Machine: *no*

Human: *can a cat fly ?*

Machine: *no .*

Human: *how many legs does a cat have ?*

Machine: *four , i think .*

Human: *how many legs does a spider have ?*

Machine: *three , i think .*

Human: *how many legs does a centipede have ?*

Machine: *eight .*

Human: *what is the color of the sky ?*

Machine: *blue .*

- Conversation 2: Job and Personality

Human: *what is your job ?*

Machine: *i 'm a lawyer .*

Human: *what do you do ?*

Machine: *i 'm a doctor .*

COMPARISON

Human: What is the deepest spot on the world ?

NCM: The pacific ocean.

CB: What ?

Human: Do you like Mexican food or Indian food ?

NCM: I hate it.

CB: I love Mexican food.

Human: Who are you crazy about ?

NCM: You're the one who's crazy.

CB: Who are you ?

Human: What's thirty plus forty ?

NCM: Sixty ?

CB: What should we talk about ?

- Evaluation Method:

- Evaluated using 200 distinct questions.
- Human judges compared NCM and CleverBot (CB).

- Results:

- NCM Preferred: 97 out of 200 questions.
- CleverBot Preferred: 60 out of 200 questions.
- Ties: 20 out of 200 questions.
- Disagreements: 23 out of 200 questions.

STRENGTHS AND WEAKNESSES

- Strengths:

- Flexibility Across Domains
- End-to-End Learning
- Simplicity: Easy to implement
- Long-range Correlations:
Effectively handles dependencies
in data across extensive
sequences.

- Weaknesses:

- Lack of Consistency
- Tends to generate brief replies
that lack depth.
- Reliance on Manual Inspection

CONVERSATION MODELING IN INTERACTIVE FICTION AND STORY GENERATION

- Dynamic Character Interactions: The model's conversational capabilities can be utilized to automate character interactions in Interactive Fiction, enabling them to adapt dynamically to user inputs.
- Storytelling: By training on diverse datasets, the model could generate stories or content based on the evolving preferences or actions of the player.

THANK YOU