A Conversational Neural Model

Taras Lehinevych

info@taraslehinevych.me
@lehinevych

Agenda

- Motivation
- Datasets
- Sequence-to-sequence model
- Improved sequence-to-sequence model
- Amazon Alexa Prize Competition - MILABOT
- Conclusion

Motivation

- Pass Turing test :)
- General purpose conversation
- Chatbots !!!
- Make better products

Task

Having previous dialog sentences and/or context information predict the next dialogs sentence.

$$P(reply \mid context)$$

$$Similarity(reply \mid context)$$

Dataset

Open Subtitles

- Open-domain conversation dataset
- Dialog tracking
- No speaker information
- Fantastic and weird responses

Twitter

- Clear tree structure of dialog
- Information about a speaker
- Tweets are based on links and photos
- No natural conversation flow

Ubuntu

- Dialogues (human-human)
- Partial speaker information
- A lot of links, commands, misspelled words and etc.

Lison, Pierre, and Raveesh Meena. "Automatic turn segmentation for Movie & TV subtitles." *Spoken Language Technology Workshop (SLT), 2016 IEEE*. IEEE, 2016.

Lowe, Ryan, et al. "The ubuntu dialogue corpus: A large dataset for research in unstructured multi-turn dialogue systems." *arXiv preprint arXiv:1506.08909* (2015).

A Neural Conversational Model

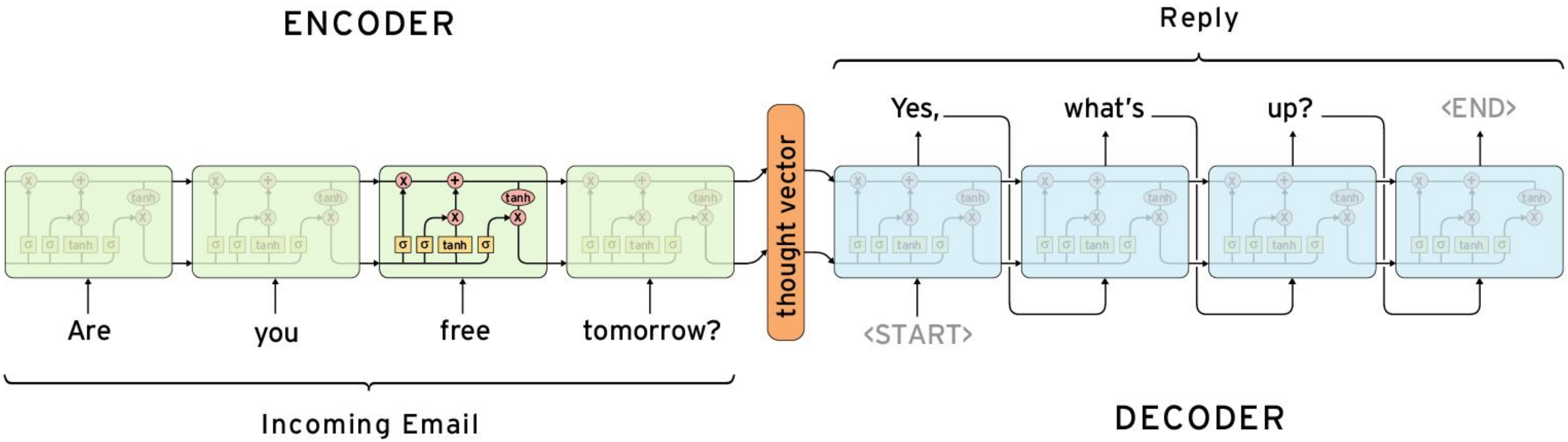


Image borrowed from [Deep Learning for Chatbots : Part 1]

[Vinyals, Oriol, and Quoc Le. "A neural conversational model." arXiv preprint arXiv:1506.05869 (2015)]

A Neural Conversational Model

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?
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 .

Human: what is the purpose of life ?
Machine: to serve the greater good .
Human: what is the purpose of living ?
Machine: to live forever .
Human: what is the purpose of existence ?
Machine: to find out what happens when we get to the planet earth .
Human: where are you now ?
Machine: i ’m in the middle of nowhere .

[Vinyals, Oriol, and Quoc Le. "A neural conversational model." arXiv preprint arXiv:1506.05869 (2015)]

A Neural Conversational Model

IT Helpdesk Troubleshooting
Common 3 problems (remote access, software crashes, password issues)

Dialog thread - 400 words
Training set - 30M tokens
Validation - 3M tokents

Common names, numbers, and full URLs - removed

1024 memory cell - LSTM
20K - vocabulary

Perplexity - 8
N-gram model perplexity - 18

OpenSubtitles
Training set - 62M sentences (923M tokens)
Validation set - 26M sentences (395M tokens)

XML tags and hyperlinks - removed

4096 memory cell - LSTM
100K - vocabulary
2048 linear unit - projection
Attention mechanism - doesn't improve perplexity

Perplexity - 17
5-gram model perplexity - 28

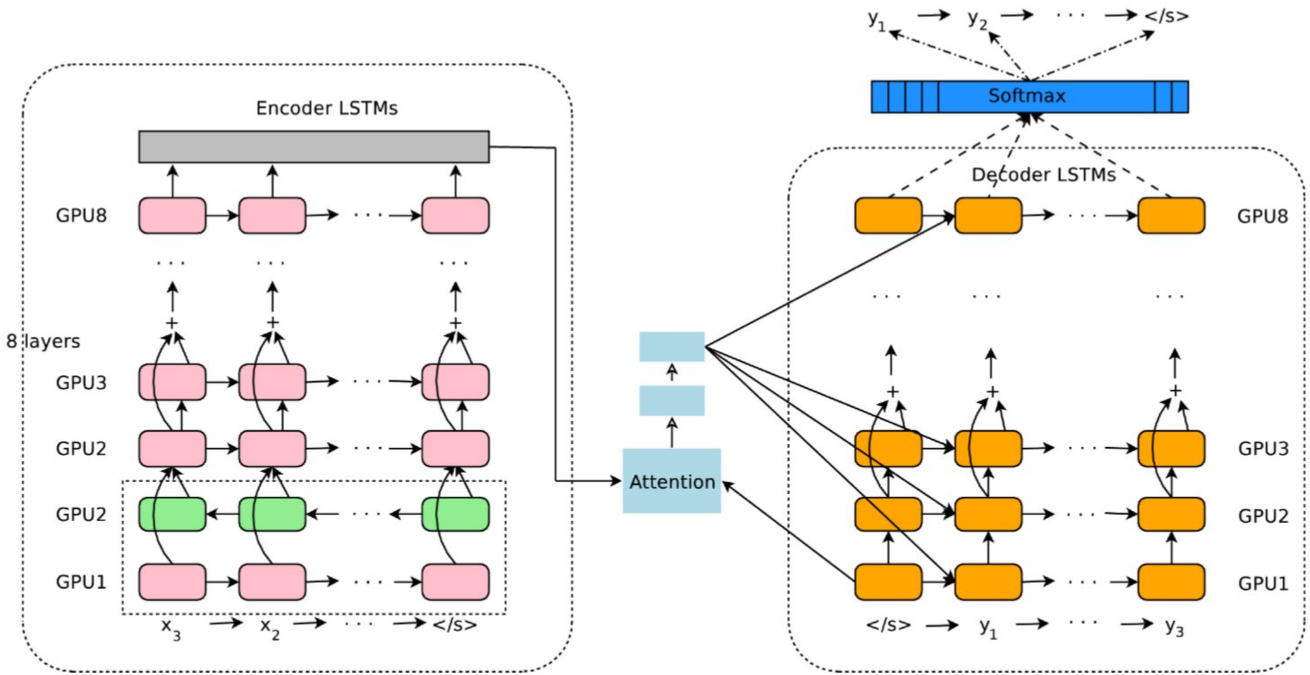
Vinyals, Oriol, and Quoc Le. "A neural conversational model." arXiv preprint arXiv:1506.05869 (2015)

Objective Function

Training maximizing given word $\log P(reply \mid context, word)$

Answer generation - maximizing given reply (beam search) $\log P(reply \mid context, word)$

Improved Seq2Seq



[Wu, Yonghui, et al. "Google's neural machine translation system: Bridging the gap between human and machine translation." *arXiv preprint arXiv:1609.08144* (2016).]

Common Answers

Human: is morality and ethics the same ?
Machine: yes , sir .
Human: what are the things that i do to be immoral?
Machine: i don ’t know .
Human: give me some examples of moral actions ...
Machine: i ’m not a moralist .
Human: alright , morality ?
Machine: integrity .
Human: what is integrity ?
Machine: i ’m sorry , i don’t know what else to say .

Input: What are you doing?			
−0.86	I don't know.	−1.09	Get out of here.
−1.03	I don't know!	−1.09	I'm going home.
−1.06	Nothing.	−1.09	Oh my god!
−1.09	Get out of the way.	−1.10	I'm talking to you.
Input: what is your name?			
−0.91	I don't know.	...	
−0.92	I don't know!	−1.55	My name is Robert.
−0.92	I don't know, sir.	−1.58	My name is John.
−0.97	Oh, my god!	−1.59	My name's John.
Input: How old are you?			
−0.79	I don't know.	...	
−1.06	I'm fine.	−1.64	Twenty-five.
−1.17	I'm all right.	−1.66	Five.
−1.17	I'm not sure.	−1.71	Eight.

Table 1: Responses generated by a 4-layer SEQ2SEQ neural model trained on 20 million conversation pairs take from the OpenSubtitles dataset. Decoding is implemented with beam size set to 200. The top examples are the responses with the highest average probability log-likelihoods in the N-best list. Lower-ranked, less-generic responses were manually chosen.

[Li, Jiwei, et al. "A diversity-promoting objective function for neural conversation models." arXiv preprint arXiv:1510.03055 (2015).]

Maximum Mutual Information (MMI)

$$\log P(reply \mid context,\ word)$$

$$\log P(reply \mid context,\ word) - a \log P(reply \mid word)$$

$$(1-a) \log P(reply \mid context,\ word) + a \log P(context \mid reply,\ word)$$

[Li, Jiwei, et al. "A diversity-promoting objective function for neural conversation models." arXiv preprint arXiv:1510.03055 (2015).]

Context Sampling

$$\log P(reply \mid context, word) - \frac{a}{k} \sum_{i}^{k} \log P(reply \mid context_i, word)$$

[Shao, Louis, et al. "Generating Long and Diverse Responses with Neural Conversation Models." *arXiv preprint arXiv:1701.03185* (2017).]

Examples

Question: What do you have planned tonight?

Seq2seq: I have no idea what I want to do with my life. I have no idea what I want to

Human ratings: Bad, Mediocre, Bad

MMI: I was planning on doing a few of these, but I’m not sure what to choose. I need to find a place to camp, and I don’t have a lot
of time to make a plan.

Human ratings: Excellent, Good, Mediocre

Examples

Question: What do you hope to see in the future?

Seq2seq: I want to be able to go back in time to when I was born. I want to be

Human ratings: Mediocre, Bad, Mediocre

MMI: I want to see an actual movie where the characters have the ability to hide the fact that they are actually in the movie

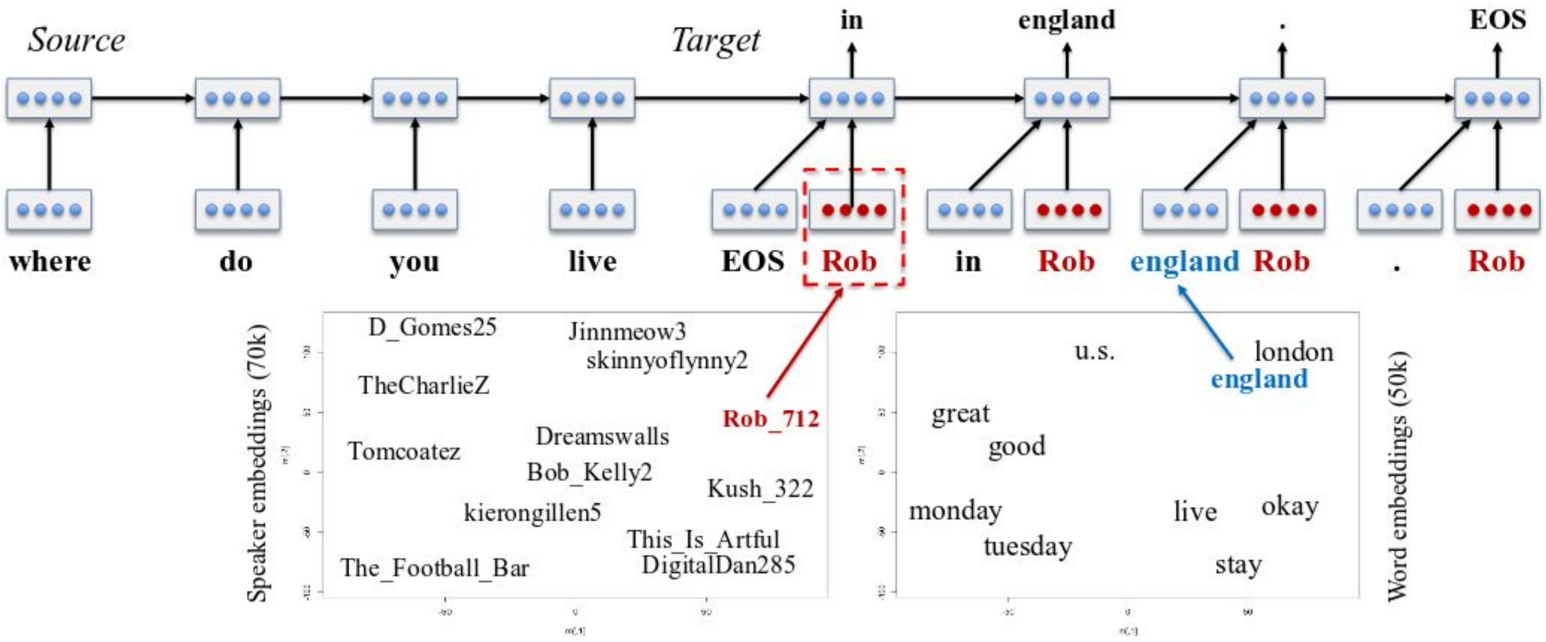
Human ratings: Good, Acceptable, Acceptable

I don’t know who am I

Human: what is your job ?
Machine: i ’m a lawyer .
Human: what do you do ?
Machine: i ’m a doctor .

Human: Where were you born?
Machine: I was born in Canada.
Human: Where are you from?
Machine: England, you?
Human: Where did you grow up?
Machine: I grew up in Texas.

A Persona-Based Neural Conversation Model



[Li, Jiwei, et al. "A persona-based neural conversation model." *arXiv preprint arXiv:1603.06155* (2016).]

Examples

Q: What company do you work for?
Baseline: I work for a company.

User1: I don't have a job.
User2: I'm a manager.
User3: I'm working. #
User4: I work for a company.
User5: I'm working. #
User6: Customer service.
User7: I work in retail.
User8: I work in a cafe.
User9: I work at a gas station.
User10: I don't know #

USER 1
Q: Where is your hometown?
A: I'm from England.

Q: Where are you from?
A: I'm from England.

Q: In which city do you live now?
A: I live in London.

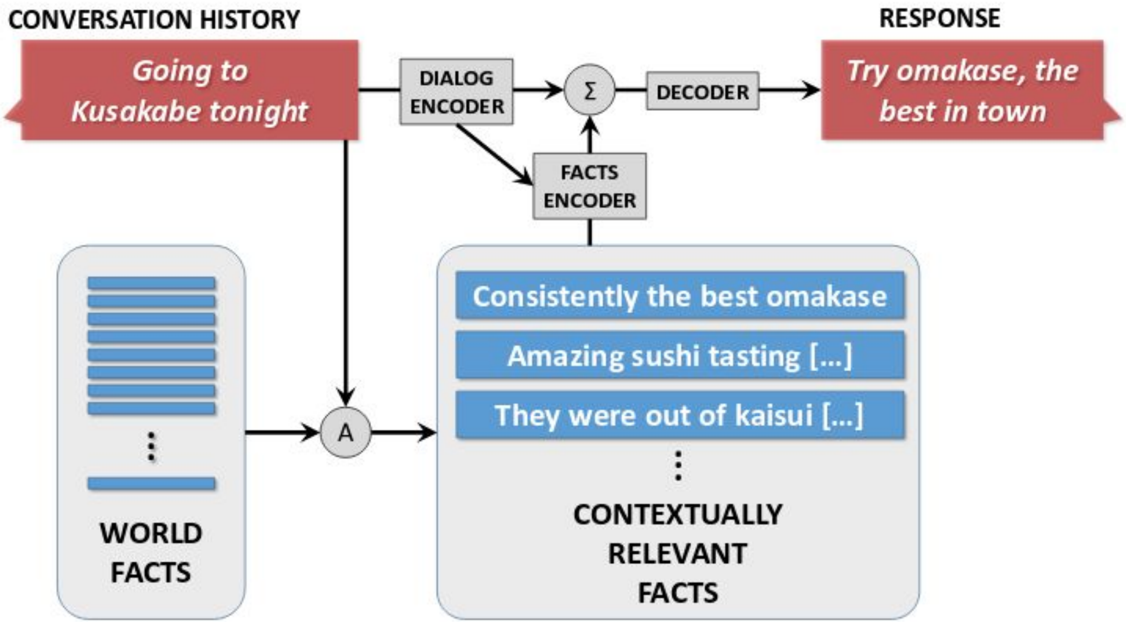
Q: In which country do you live now?
A: I live in England.

Q: What is your major?
A: Business. You?

Q: What did you study in college?
A: I did business studies.

[Li, Jiwei, et al. "A persona-based neural conversation model." *arXiv preprint arXiv:1603.06155* (2016).]

A Knowledge-Grounded Neural Conversation Model



Ghazvininejad, Marjan, et al. "A Knowledge-Grounded Neural Conversation Model." *arXiv preprint arXiv:1702.01932* (2017).

Lack of Proper Metrics

	Twitter				Ubuntu			
Metric	**Spearman**	p-value	**Pearson**	p-value	**Spearman**	p-value	**Pearson**	p-value
Greedy	0.2119	0.034	0.1994	0.047	0.05276	0.6	0.02049	0.84
Average	0.2259	0.024	0.1971	0.049	-0.1387	0.17	-0.1631	0.10
Extrema	0.2103	0.036	0.1842	0.067	0.09243	0.36	-0.002903	0.98
METEOR	0.1887	0.06	0.1927	0.055	0.06314	0.53	0.1419	0.16
BLEU-1	0.1665	0.098	0.1288	0.2	-0.02552	0.8	0.01929	0.85
BLEU-2	0.3576	< 0.01	0.3874	< 0.01	0.03819	0.71	0.0586	0.56
BLEU-3	0.3423	< 0.01	0.1443	0.15	0.0878	0.38	0.1116	0.27
BLEU-4	0.3417	< 0.01	0.1392	0.17	0.1218	0.23	0.1132	0.26
ROUGE	0.1235	0.22	0.09714	0.34	0.05405	0.5933	0.06401	0.53
Human	0.9476	< 0.01	1.0	0.0	0.9550	< 0.01	1.0	0.0

Table 3: Correlation between each metric and human judgements for each response. Correlations shown in the human row result from randomly dividing human judges into two groups.

[Liu, Chia-Wei, et al. "How NOT to evaluate your dialogue system: An empirical study of unsupervised evaluation metrics for dialogue response generation." *arXiv preprint arXiv:1603.08023* (2016).]

Amazon Alexa Prize competition

Allow 200,000 labels on Amazon Mechanical Turk

Maintain over 32 dedicated Tesla K80 GPUs for running our live system.

Average user score of 3.15 on a scale 1−5
Minimal number of hand-crafted states and rules and without engaging in non-conversational activities (such as playing games or taking quizzes).

System averaged a high 14.5 −16.0 turns per dialogue

MILABOT is likely to be the most engaging system among all systems in the competition.

All system components are learnable.

[Serban, Iulian V., et al. "A deep reinforcement learning chatbot." *arXiv preprint arXiv:1709.02349* (2017).]

Ensemble of 22 models

Template-based models: Alicebot, Elizabot, Initiatorbot, Storybot

Knowledge Base-based Question Answering: Evibot, BoWMovies

Retrieval-based Neural Networks: VHRED models (VHREDRedditPolitics, VHREDRedditNews, VHREDRedditSports, VHREDRedditMovies, VHREDWashingtonPost, VHREDSubtitles), SkipThought Vector Models, Dual Encoder Models, Bag-of-words Retrieval Models

Retrieval-based Logistic Regression: BoWEscapePlan,

Search Engine-based Neural Networks: LSTMClassifierMSMarco

Generation-based Neural Networks: GRUQuestionGenerator

Serban, Iulian V., et al. "A deep reinforcement learning chatbot." *arXiv preprint arXiv:1709.02349* (2017).

MILABOT

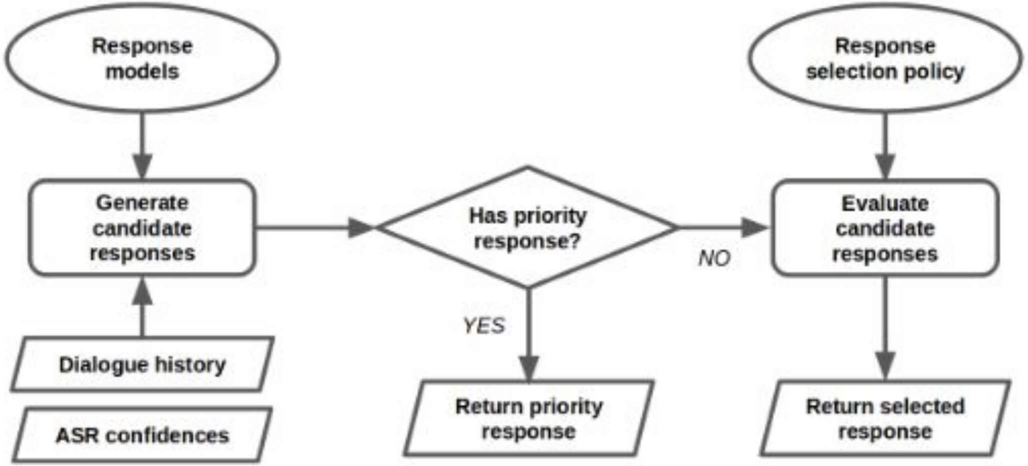


Figure 1: Dialogue manager control flow.

[Serban, Iulian V., et al. "A deep reinforcement learning chatbot." *arXiv preprint arXiv:1709.02349* (2017).]

Scoring Model

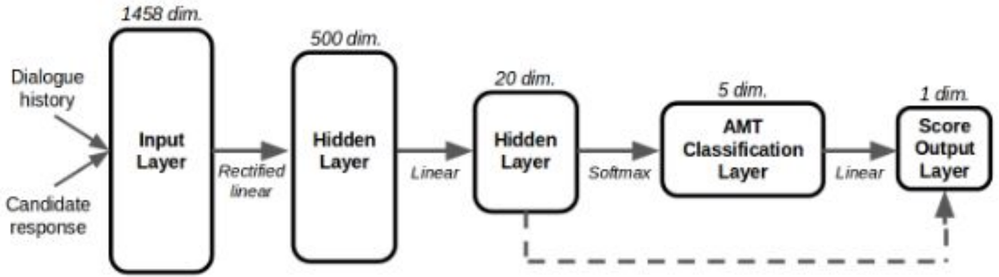


Figure 2: Computational graph for scoring model, used for the model selection policies based on both action-value function and stochastic policy parametrizations. The model consists of an input layer with 1458 features, a hidden layer with 500 hidden units, a hidden layer with 20 hidden units, a softmax layer with 5 output probabilities (corresponding to the five AMT labels in Section 4.3), and a scalar-valued output layer. The dashed arrow indicates a skip connection.

Serban, Iulian V., et al. "A deep reinforcement learning chatbot." *arXiv preprint arXiv:1709.02349* (2017).

Response Evaluation

Trade-off between immediate and long-term user satisfaction

Reward = expected cumulative return

$$R = \sum_{t=1}^{T} \gamma^{t} r_{t}$$

[Serban, Iulian V., et al. "A deep reinforcement learning chatbot." *arXiv preprint arXiv:1709.02349* (2017).]

Conclusions

Reading List

Generative Deep Neural Networks for Dialogue: A Short Review

A Survey on Dialogue Systems: Recent Advances and New Frontiers

A Deep Reinforcement Learning Chatbot

Thanks. Questions?