Integrating Chatbots With Other Knowledge

Head First Theory and Practice

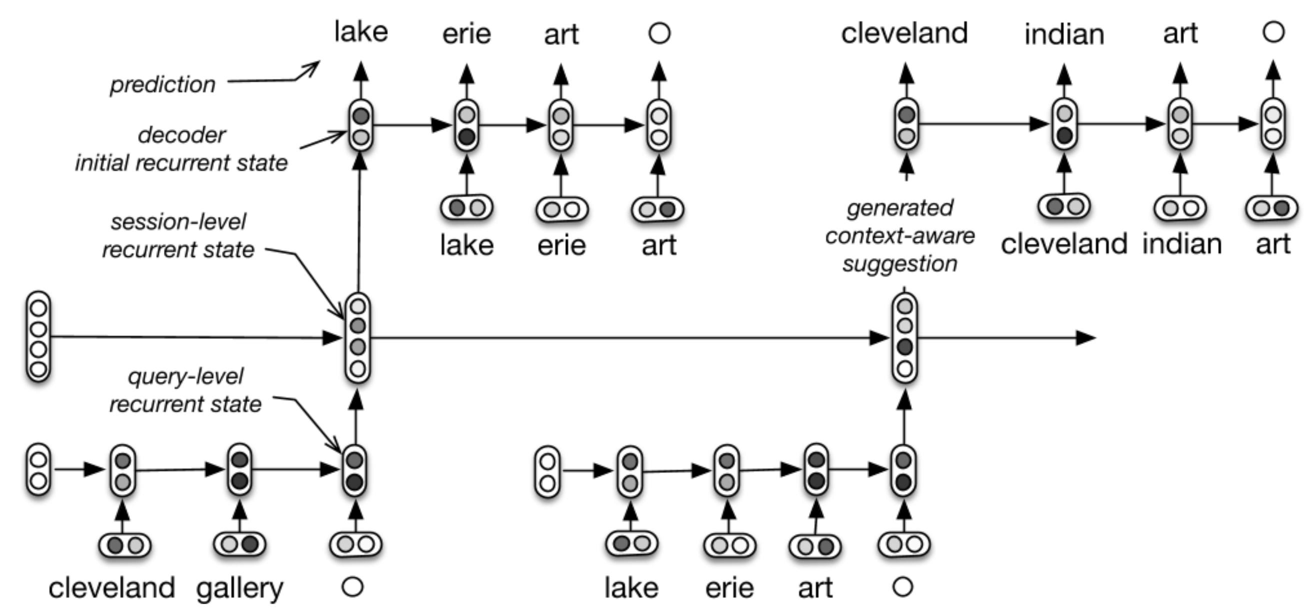
Yanran Li
The Hong Kong Polytechnic University

Key Issues

And Future Directions

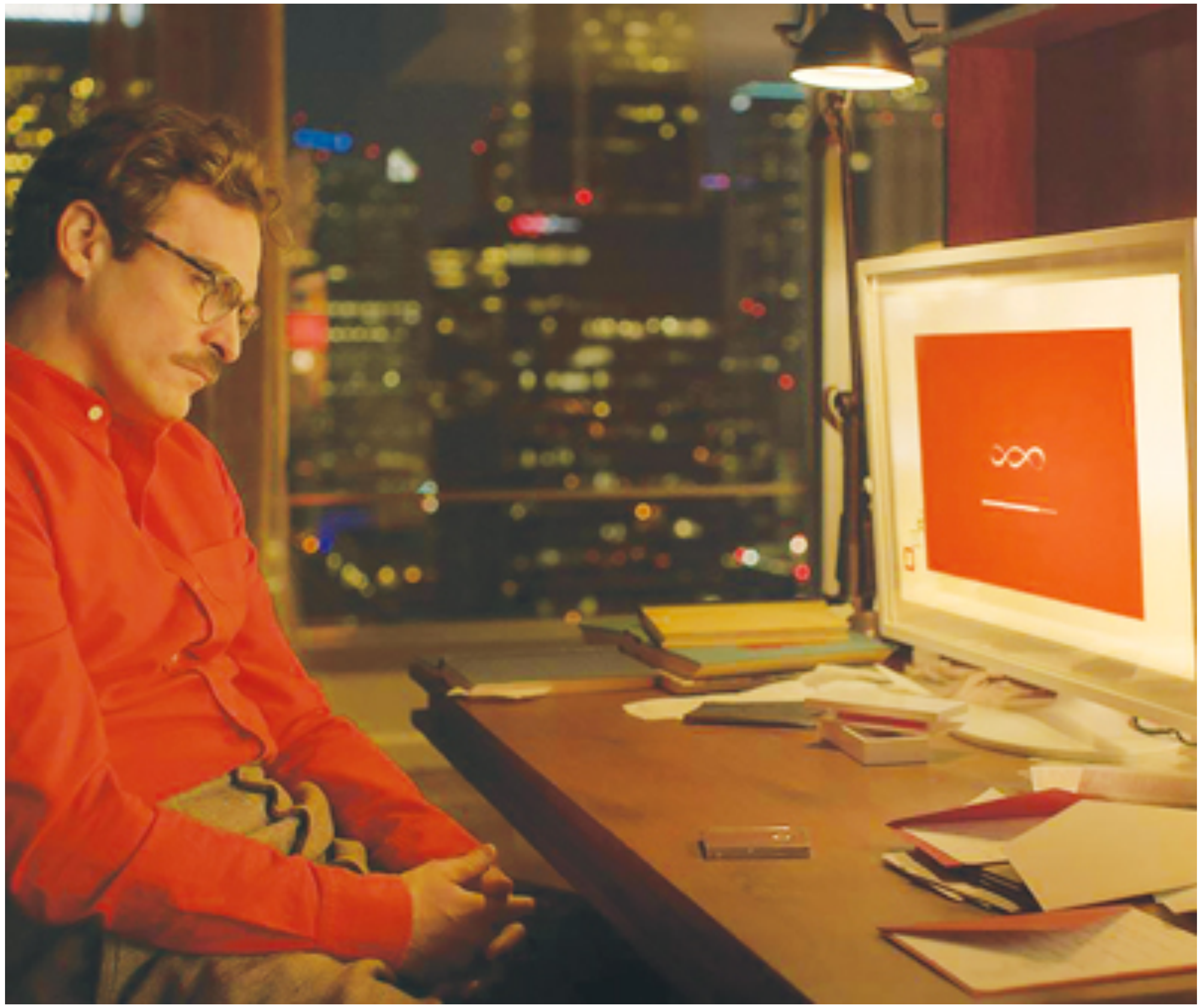
Memory

- Context, Dialog History
  - Hierarchical Recurrent Encoder-Decoder[18]
  - (Variable) HRED[19]



Human-like, Careful, Friendly

- Emotional
  - “Her” in the science fiction film
  - “smart-aging”

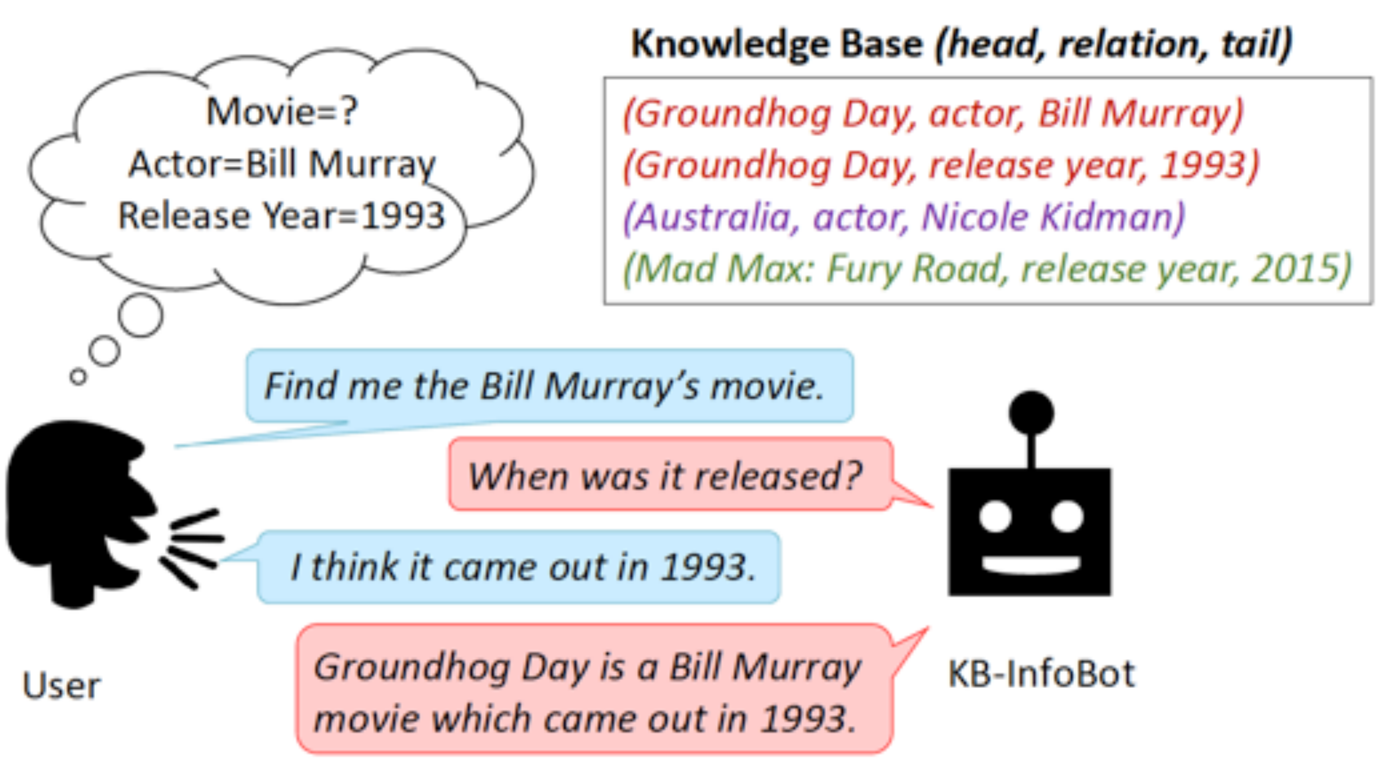


Knowledgable

- Commonsense Knowledge
- Other Knowledge

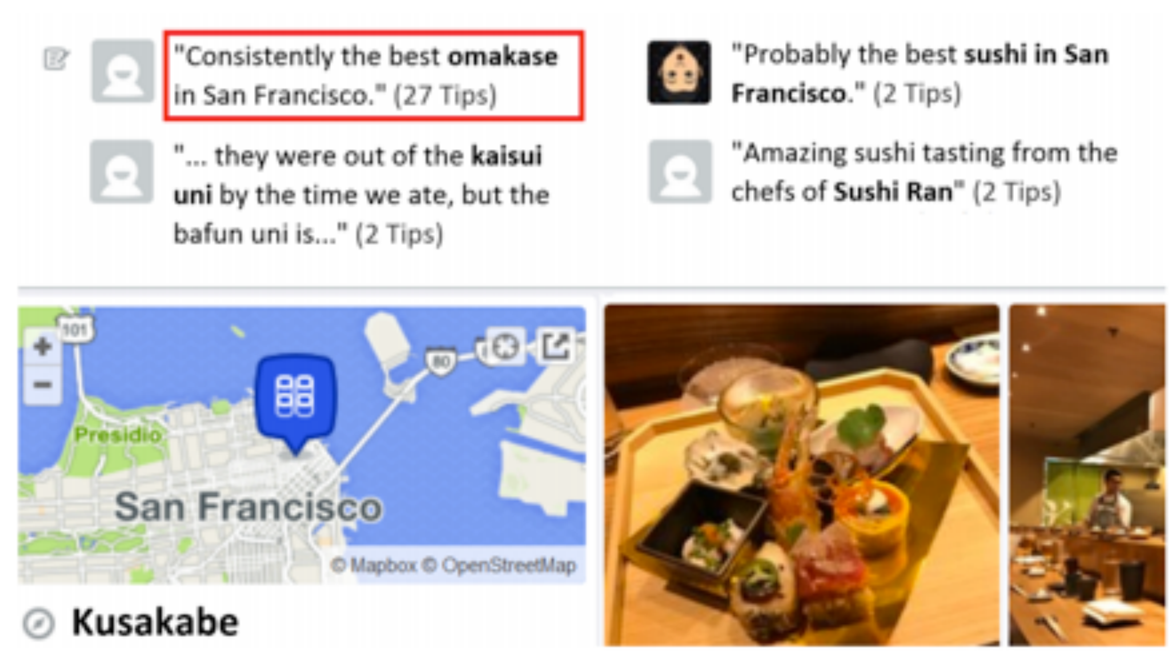
Commonsense Knowledge Augumentation

- Fact
- Entity
- *Linguistic knowledge
- Knowledge base and knowledge graph



Other Knowledge Augumentation

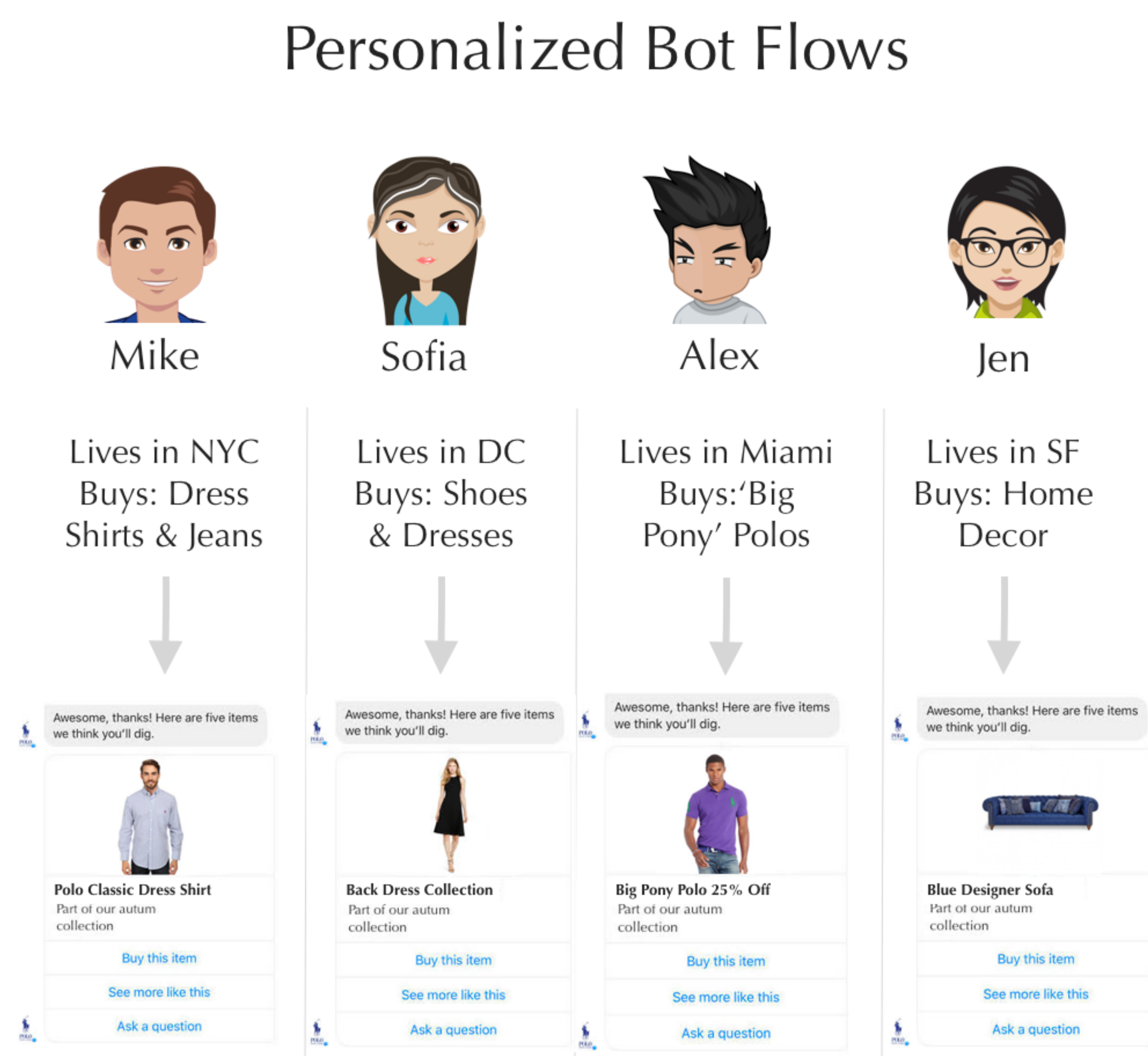
- Topic
- History
- Scenario (e.g. image)
- News/Articles/Rationales

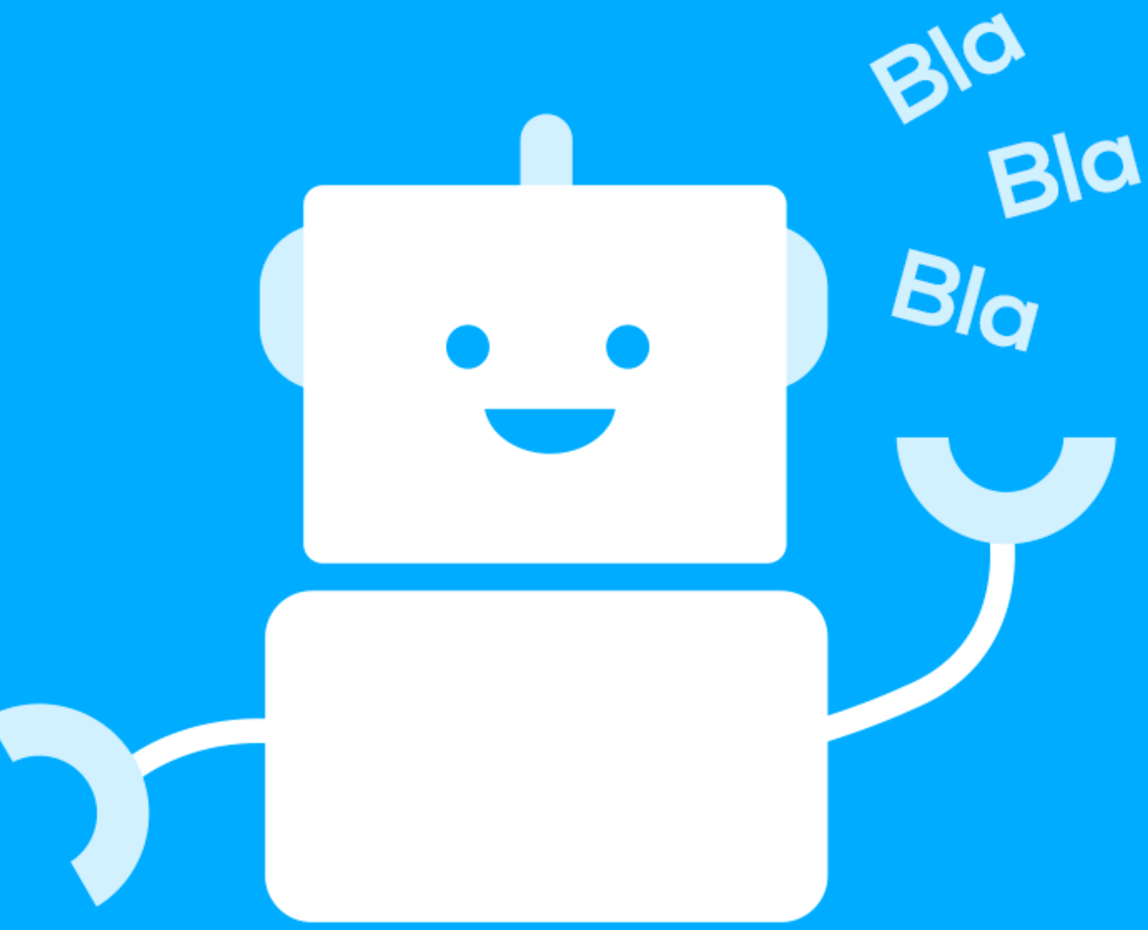


User input: *Going to Kusakabe tonight.*
Neural model: Have a great time!
Human: You'll love it! Try omasake, the best in town.

Personalized

- Transfer Learning
- Lifelong Learning
- *Lack of Personal Data





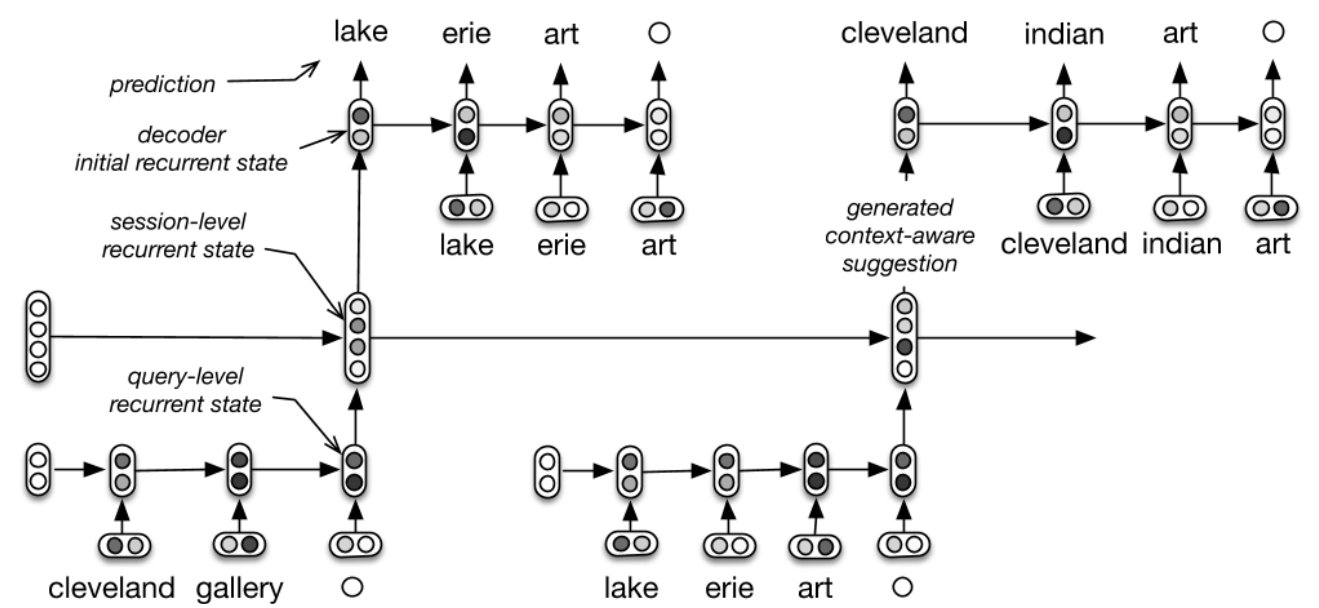
VS.

Context, History

Conversational Agents

Memory

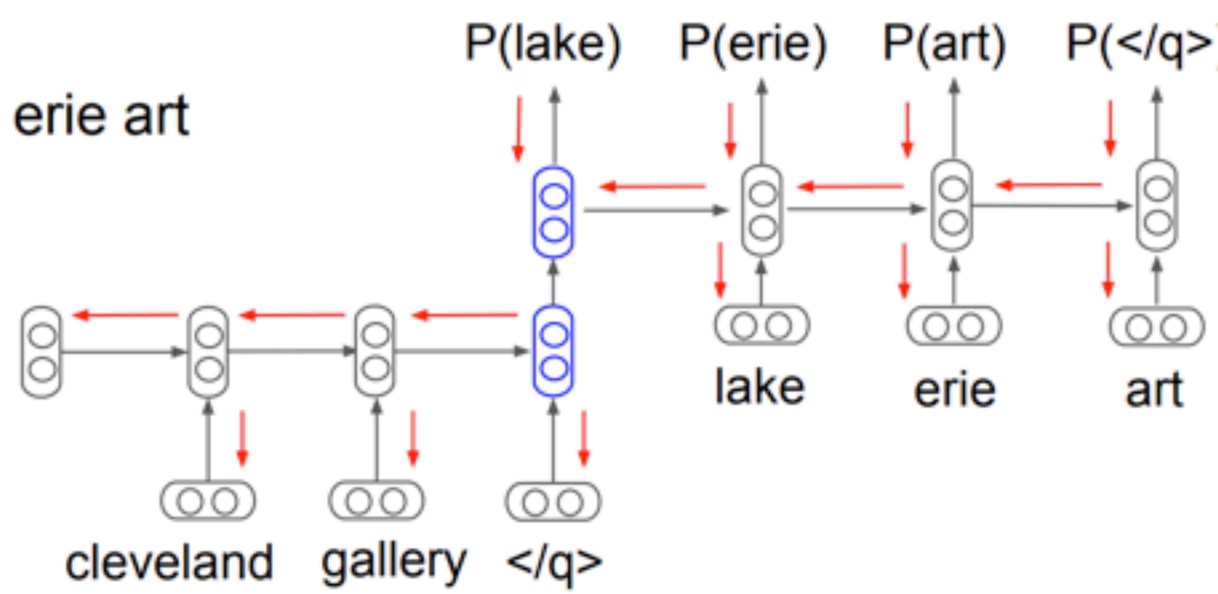
- Context, Dialog History
  - Hierarchical Recurrent Encoder-Decoder[18]
  - (Variable) HRED[19]



Recurrent Encoder-Decoder (RED)

- A RNN encoder-decoder (RED) learns a probability distribution over the nextquery in the session given the previous one. [20]

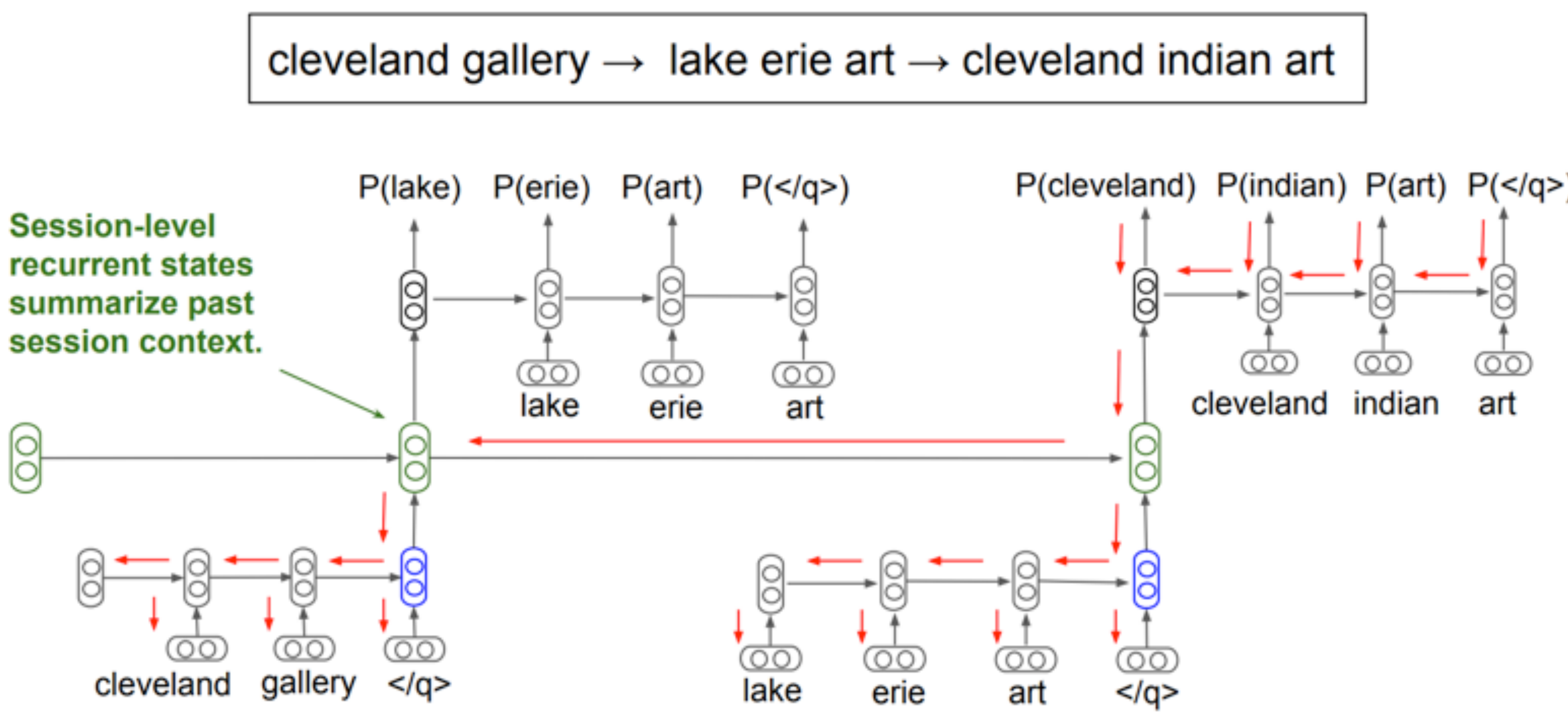
S = cleveland gallery → lake erie art



Backprop Training: $L = \log P(Q_{t+1}|Q_t) = \sum_{w_n \in Q_{t+1}} \log P(w_n|w_{<n}, Q_t)$

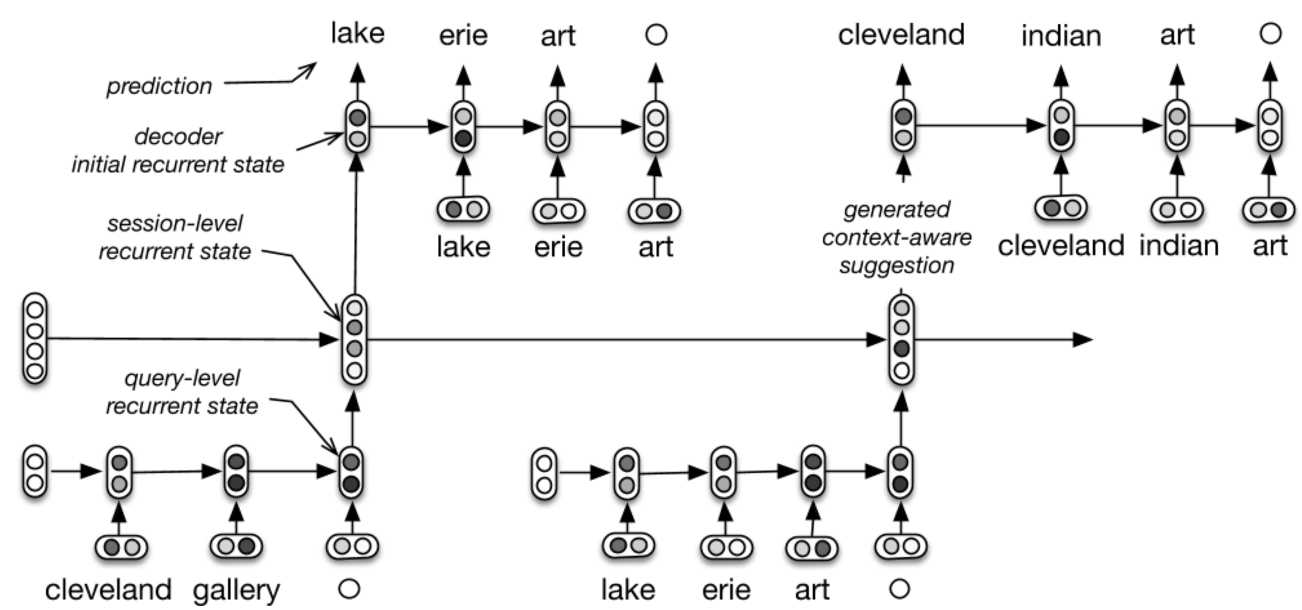
Hierarchical Recurrent Encoder-Decoder (HRED)

- The dialogue as a two-level system: a sequence of utterances, each of which is in turn a sequence of words. To model this two-level system, HRED Use an additional RNN to model the sequences of utterances in a session. [20]

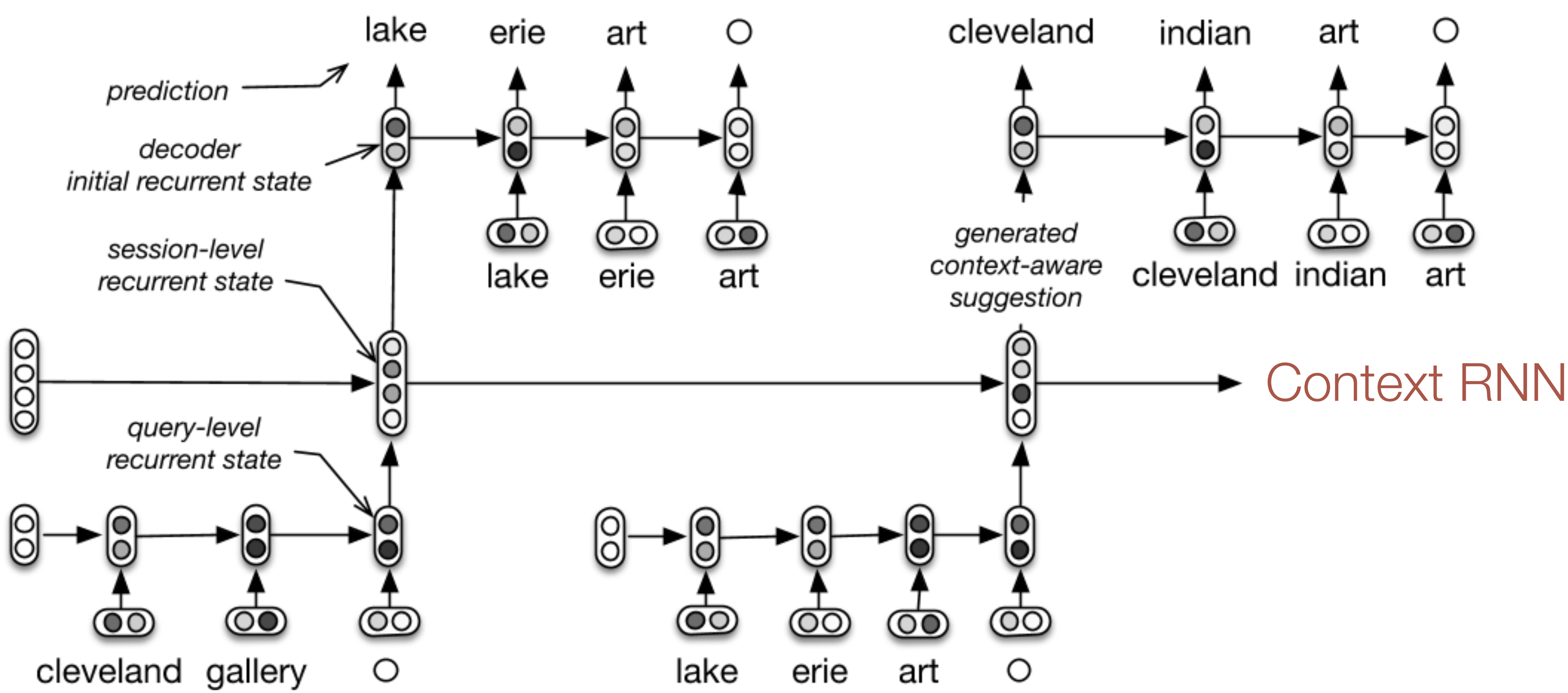


Hierarchical Recurrent Encoder-Decoder (HRED)

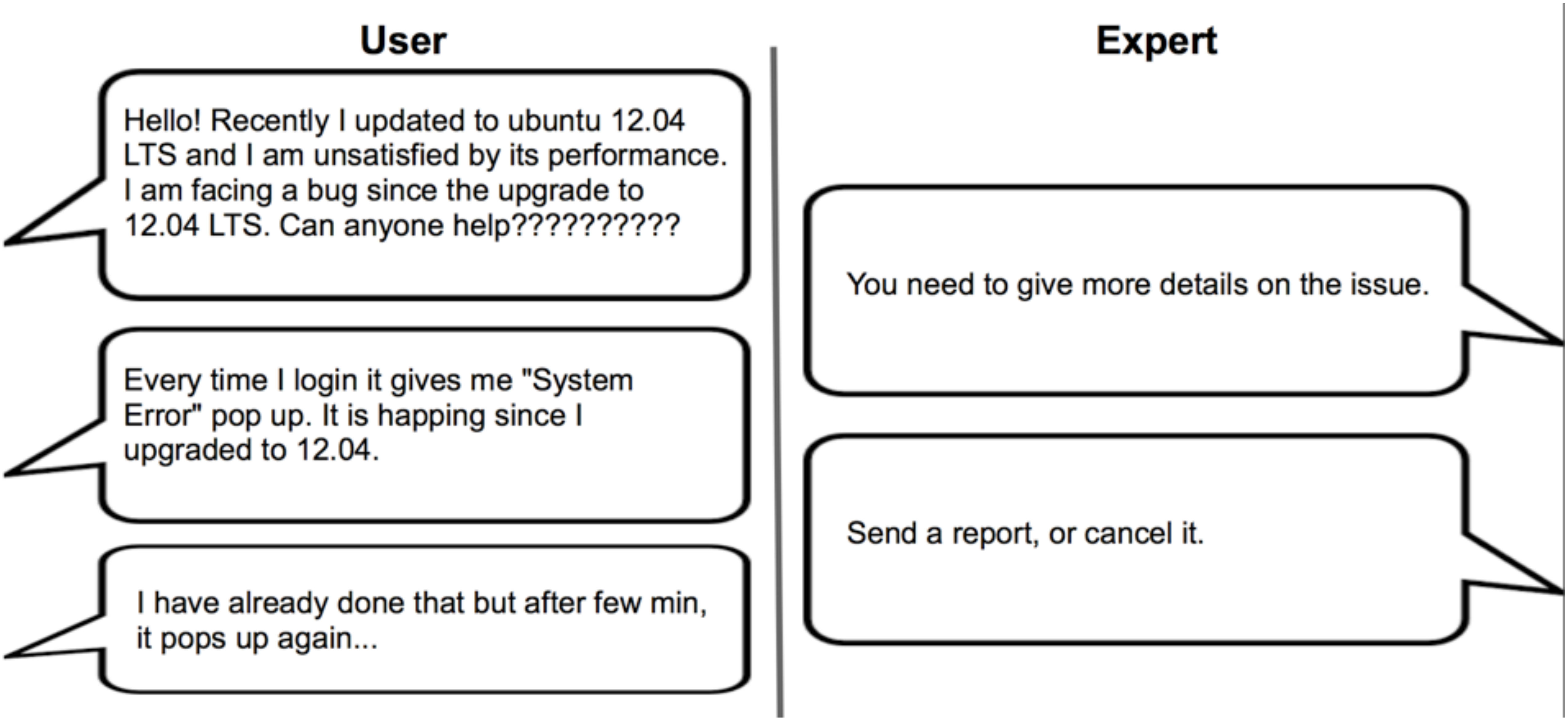
- Encoder RNN
  - encoding each utterance independently into an utterance vector
- Context RNN
  - encoding the topic/context of the dialogue up till the current utterance using utterance vectors
- Decoder RNN
  - predicting the next utterance



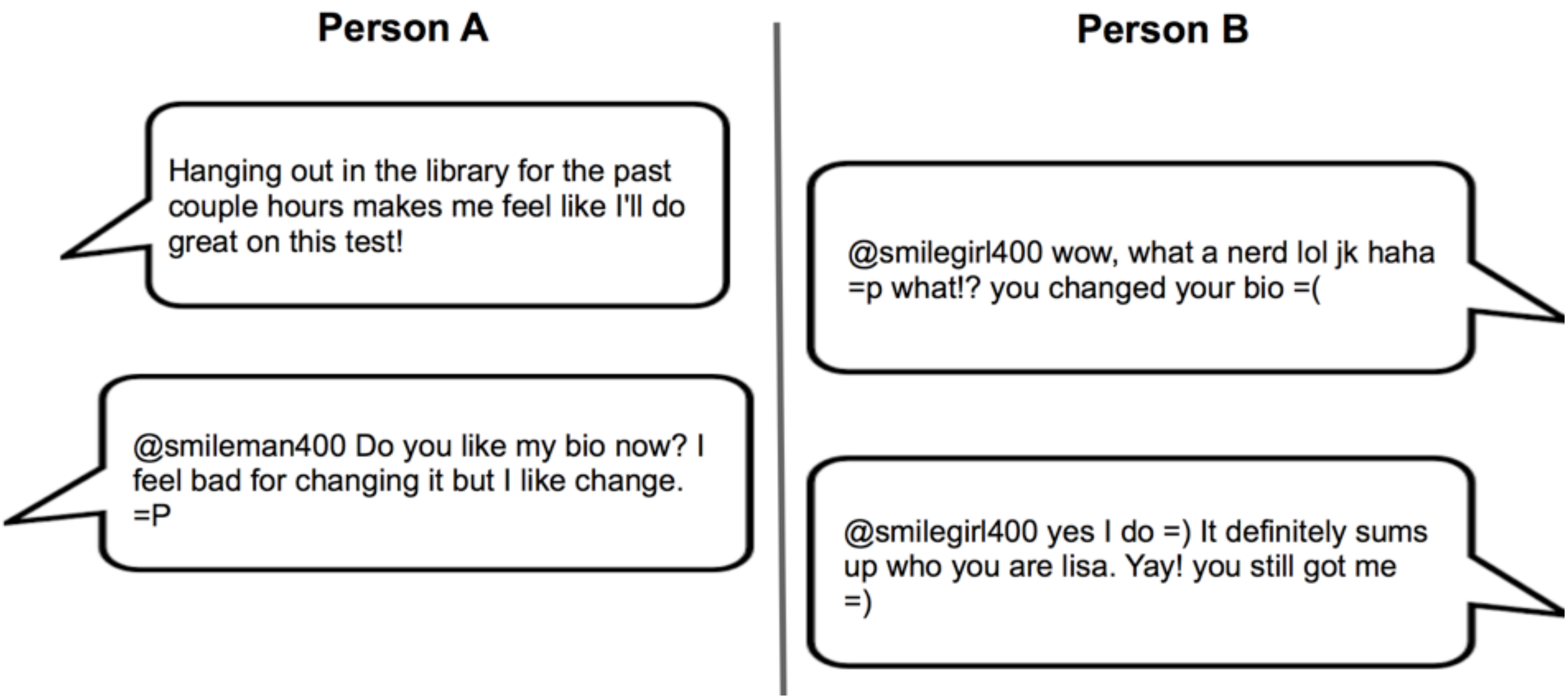
Hierarchical Recurrent Encoder-Decoder (HRED)

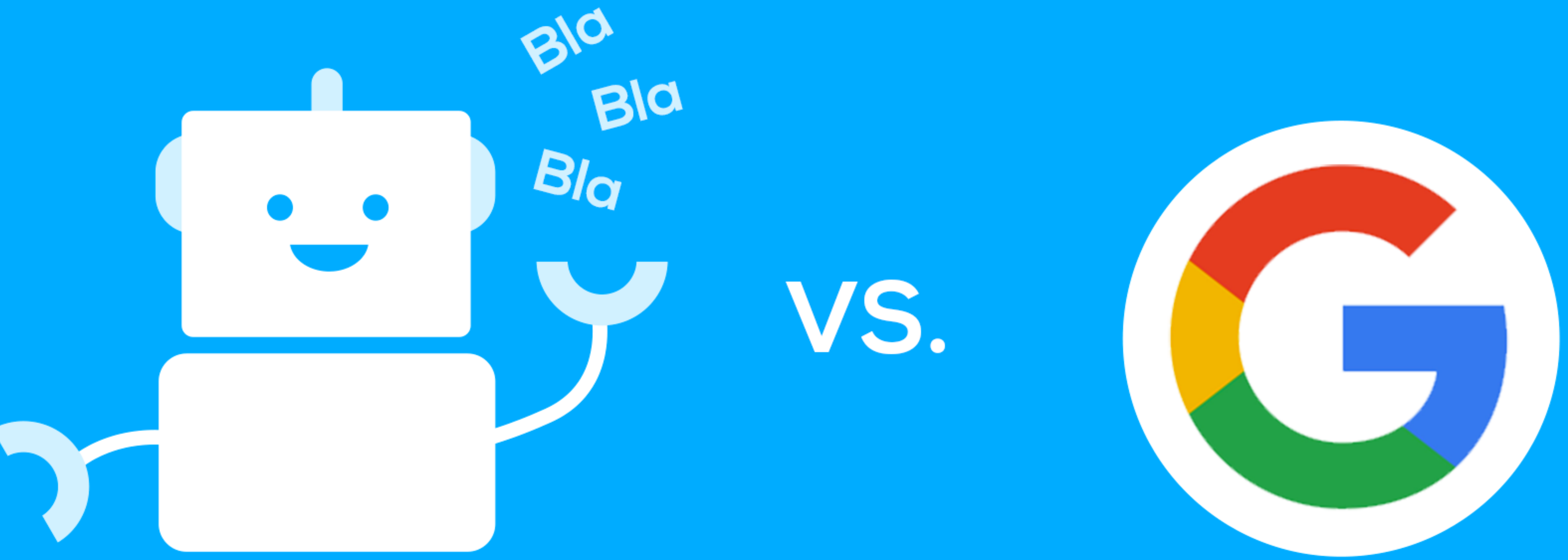


Example - Ubuntu Corpus



Example - Twitter Corpus





Emotional | Conversational Agents

Emotional Response

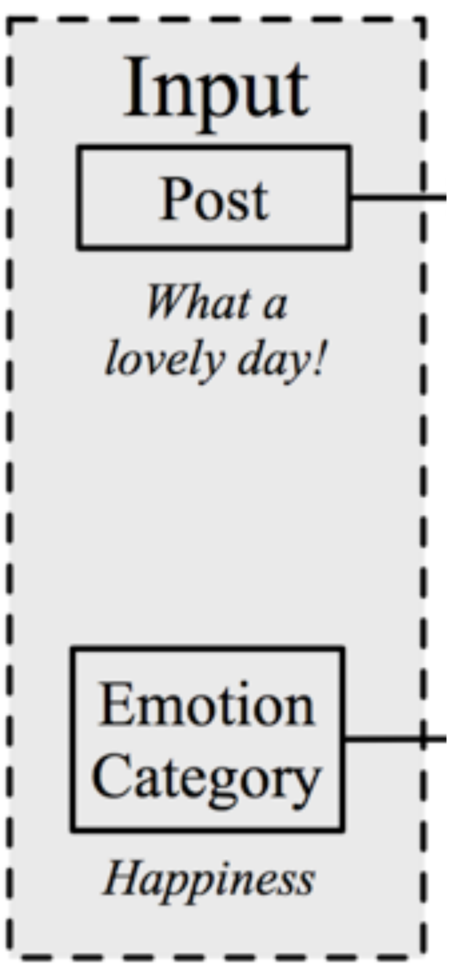
- Emotional Chatting Machine [21]

User: Worst day ever. I arrived late because of the traffic.
Chatbot (Basic Seq2Seq): You were late.
ECM (*Like*): I am always here to support you.
ECM (*Happiness*): Keep smiling! Things will get better.
ECM (*Sadness*): It's depressing.
ECM (*Disgust*): Sometimes life just sucks.
ECM (*Anger*): The traffic is too bad!

Table 1: Example conversations with/without Emotional Intelligence.

Emotional Response

- Emotional Chatting Machine [21]



Training	**posts**	217,905	
	responses	*Anger*	234,635
		Disgust	689,295
		Happiness	306,364
		Like	1,226,954
		Sadness	537,028
		Other	1,365,371
Validation	**posts**	1,000	
Test	**posts**	1,000	

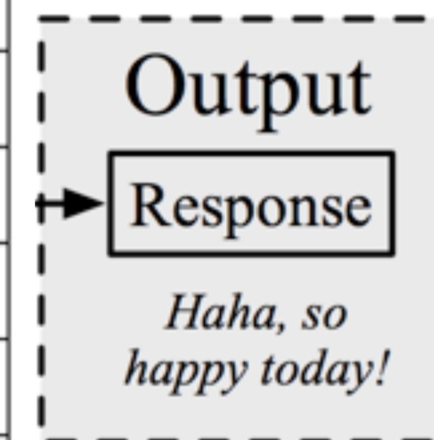


Table 3: Statistics of the Emotionalized STC Dataset.

Emotional Classifier

- Emotional Chatting Machine [21]

Training	**posts**	217,905	
	responses	*Anger*	234,635
		Disgust	689,295
		Happiness	306,364
		Like	1,226,954
		Sadness	537,028
		Other	1,365,371
Validation	**posts**	1,000	
Test	**posts**	1,000	

Table 3: Statistics of the Emotionalized STC Dataset.

Method	Accuracy
Dictionary	0.432
RNN	0.564
LSTM	0.594
Bi-LSTM	0.623

Table 2: The accuracy of emotion classifiers on th
NLPCC dataset.

Example

- Emotional Chatting Machine [21]

Post	Model		Response(Original)	Response(Translated)
比情人节更庸俗的事就是，对情人节的接待规格待遇提出要求。 What is more vulgar than Valentine's day is to make specific demands on the reception of the Valentine's day.	Seq2Seq		这才是真正的情人节。	This is a real Valentine's day.
	ECM	*Like*	情人节快乐！	Happy Valentine's Day!
		Happiness	哈哈，这也太浪漫了吧！	Aha, this is too romantic!
		Sadness	我也想要这样的情人节。	I also want this kind of Valentine's day, sigh.
		Disgust	这就是所谓的情人节！	This is the so-called Valentine's day!
		Anger	这是赤裸裸的炫耀啊！	This is shameless show-off!
知道真相的我眼泪笑下来。 Knowing the truth, I laughed myself to tears.	Seq2Seq		知道真相的我眼泪掉下来。	After knowing the truth, I burst into tears.
	ECM	*Like*	太可爱了，太可爱了！	Too lovely, it’s too lovely.
		Happiness	笑得我眼泪都出来了。	I can’t help laugh myself to tears.
		Sadness	知道真相的我眼泪掉下来。	After knowing the truth, I burst into tears.
		Disgust	这是真的还是假的？	Is this true?
		Anger	什么意思？没看懂。	What do you mean? I don’t understand.

Figure 4: Sample responses generated by Seq2Seq and ECM (original Chinese and English translation). The corresponding posts did not appear in the training set.

Example

- Emotional Chatting Machine [21]

Post	Model		Response(Original)	Response(Translated)
比情人节更庸俗的事就是，对情人节的接待规格待遇提出要求。 What is more vulgar than Valentine's day is to make specific demands on the reception of the Valentine's day.	Seq2Seq		这才是真正的情人节。	This is a real Valentine's day.
	ECM	*Like*	情人节快乐！	Happy Valentine's Day!
		Happiness	哈哈，这也太浪漫了吧！	Aha, this is too romantic!
		Sadness	我也想要这样的情人节。	I also want this kind of Valentine's day, sigh.
		Disgust	这就是所谓的情人节！	This is the so-called Valentine's day!
		Anger	这是赤裸裸的炫耀啊！	This is shameless show-off!
知道真相的我眼泪笑下来。 Knowing the truth, I laughed myself to tears.	Seq2Seq		知道真相的我眼泪掉下来。	After knowing the truth, I burst into tears.
	ECM	*Like*	太可爱了，太可爱了！	Too lovely, it’s too lovely.
		Happiness	笑得我眼泪都出来了。	I can’t help laugh myself to tears.
		Sadness	知道真相的我眼泪掉下来。	After knowing the truth, I burst into tears.
		Disgust	这是真的还是假的？	Is this true?
		Anger	什么意思？没看懂。	What do you mean? I don’t understand.

Figure 4: Sample responses generated by Seq2Seq and ECM (original Chinese and English translation). The corresponding posts did not appear in the training set.

Conditional Emotional Generation

- Incorporate Variable using Variational Approaches [19, 22]

Latent Variable Hierarchical Recurrent Encoder-Decoder (VHRED)

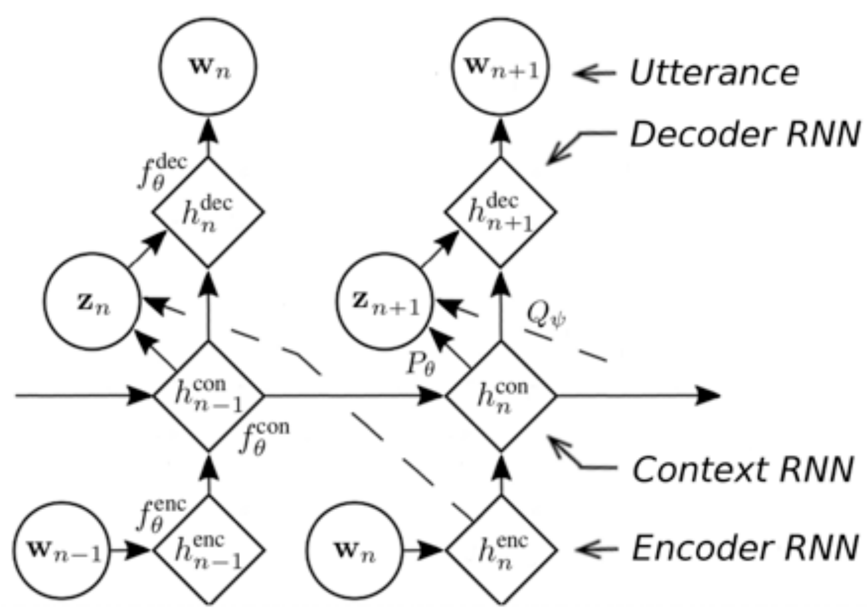


Figure 1: VHRED computational graph. Diamond boxes represent deterministic variables and rounded boxes represent stochastic variables. Full lines represent the generative model and dashed lines represent the approximate posterior model.

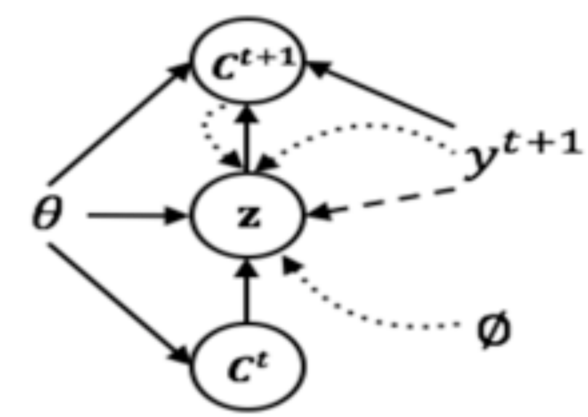


Figure 2: Graphical model for the conditional variational framework. Solid lines denote generative model $P_\theta(\mathbf{z}_n|\mathbf{y}_n, \mathbf{w}_1^{n-1})$ and $P_\theta(\mathbf{w}_n \mid \mathbf{y}_n, \mathbf{z}_n, \mathbf{w}_1^{n-1})$. When y^{t+1} is known, there exists an additional link from y^{t+1} to z (dashed line). C^t encodes context information up to time t. Dotted lines are posterior approximation $Q_\phi(\mathbf{z}_n|\mathbf{y}_n, \mathbf{w}_1^n)$.

Conditional Emotional Generation

- Incorporate Variable using Variational Approaches [19, 22]

Latent Variable Hierarchical Recurrent Encoder-Decoder (VHRED)

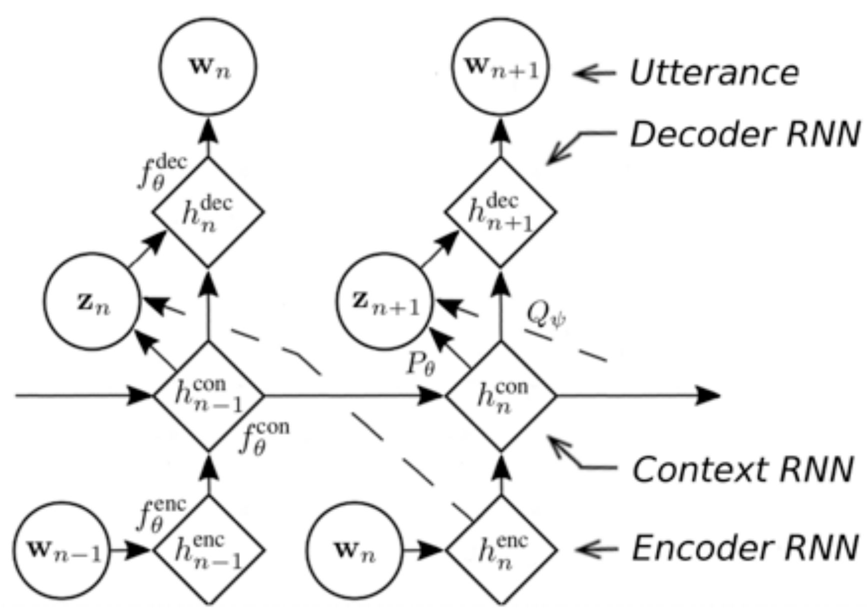


Figure 1: VHRED computational graph. Diamond boxes represent deterministic variables and rounded boxes represent stochastic variables. Full lines represent the generative model and dashed lines represent the approximate posterior model.

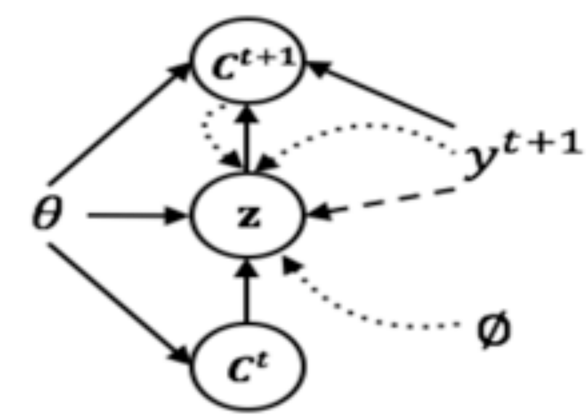


Figure 2: Graphical model for the conditional variational framework. Solid lines denote generative model $P_\theta(\mathbf{z}_n|\mathbf{y}_n, \mathbf{w}_1^{n-1})$ and $P_\theta(\mathbf{w}_n \mid \mathbf{y}_n, \mathbf{z}_n, \mathbf{w}_1^{n-1})$. When y^{t+1} is known, there exists an additional link from y^{t+1} to z (dashed line). C^t encodes context information up to time t. Dotted lines are posterior approximation $Q_\phi(\mathbf{z}_n|\mathbf{y}_n, \mathbf{w}_1^n)$.

Be Careful

- Controlling Linguistic Aspects [23]

Parameter	Value
Theme	Other
Sentiment	Negative
Professional	False
Personal	False
Length	11-20 words
Descriptive	True

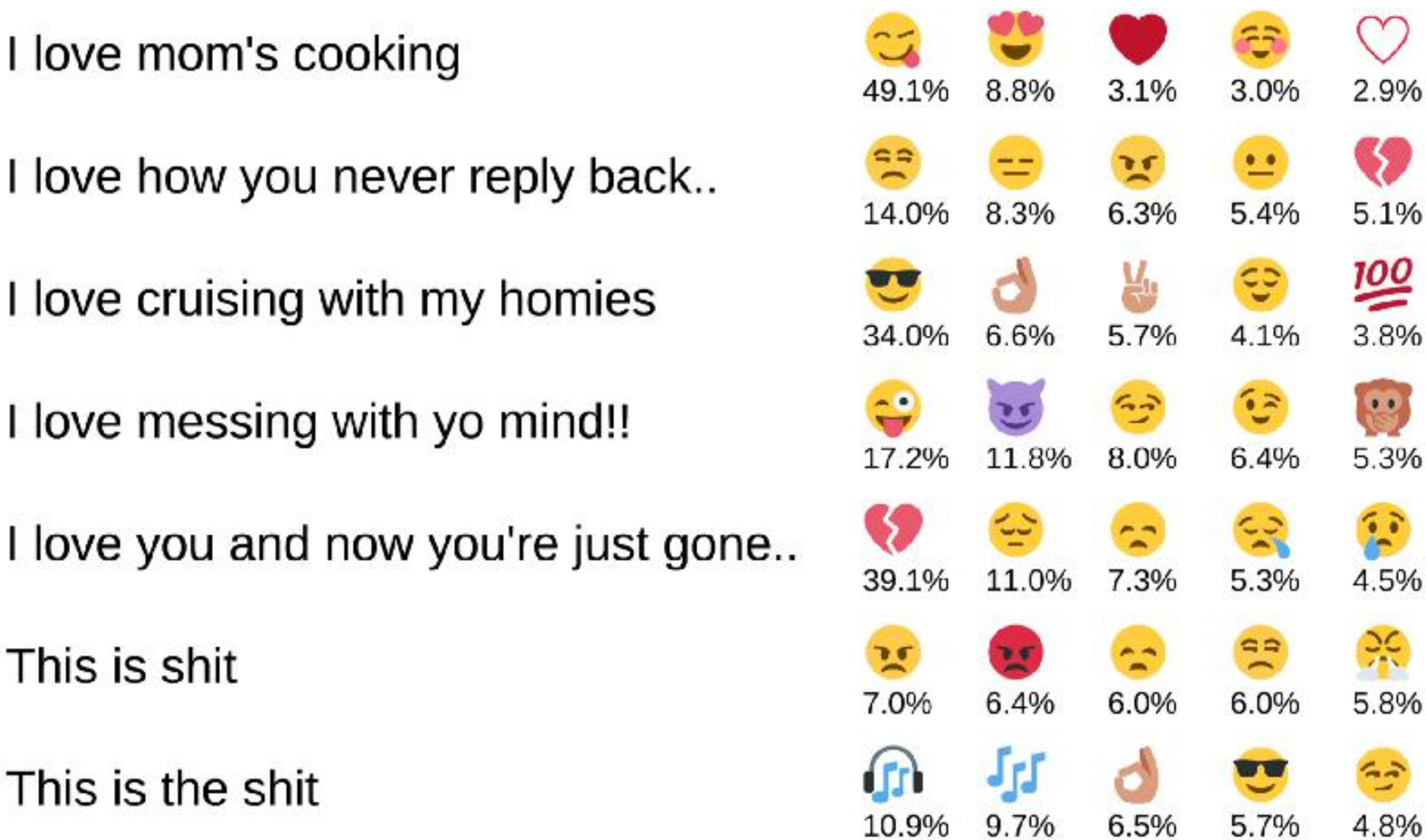
- “A little bit of a predictable and boring romantic comedy with a few funny moments but overall pretty entertaining.”
- “With such a great premise ,Escape From Tomorrow is pretty damn terrible, horrible, and no exception.”
- “The first half is lazy and stupid, but there’s a handful of funny moments which are pretty decent.”
- “There’s no denying the fact that this movie is such a horrible movie with a few bad moments.”
- “The last part of the movie just let me down, but the whole thing is pretty good.”
- “It is a little difficult to follow, but this is a rare right choice for the respective aspects of film.”
- “My biggest issue is that the first half is pretty boring, plodding, and too obvious to be honest.”

Parameter	Value
Theme	Other
Sentiment	Positive
Professional	True
Personal	False
Length	11-20 words
Descriptive	False

- “This is a must see for fans of Bergman’s “ American Dream ”, “ The Roots ”.”
- “The film’s ultimate pleasure if you want to fall in love with the ending, you won’t be disappointed”
- “One of the most inspirational films I’ve seen in years, hence the most influential, the best.”
- “The monetary system is a bit too intelligent and at times implanted in the middle of the film.”
- “The film doesn’t seem to be anything more than a dozen other films with a sophisticated vigor.”
- “The film’s late 19th century Denmark’s history is a feast for the eyes and the laughter on the screen.”
- “ The film’s simple, and a refreshing take on the complex family drama of the regions of human intelligence.”

Still some fun

- Predict Emoji [24]

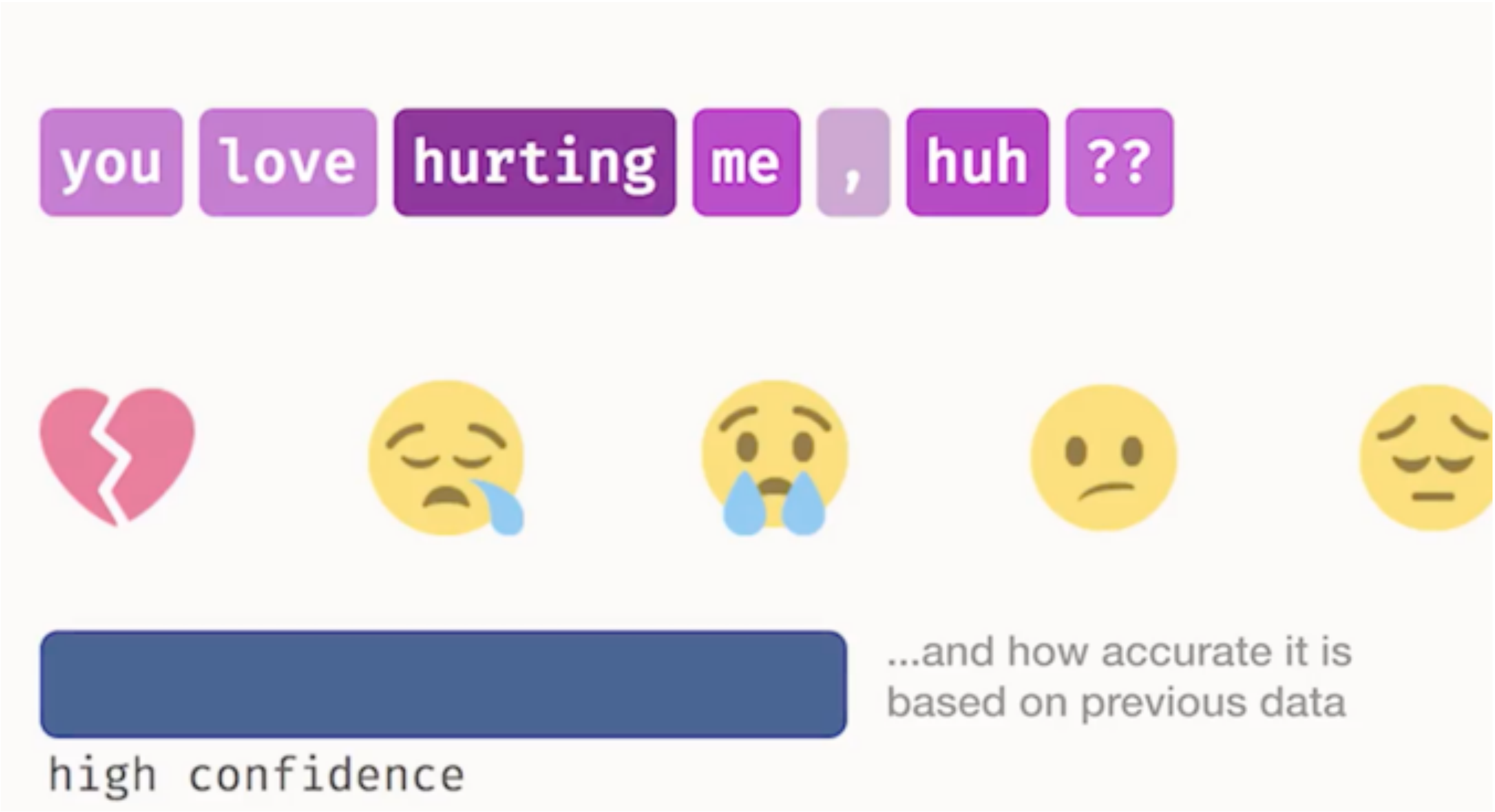


Still some fun

- Predict Emoji [24]

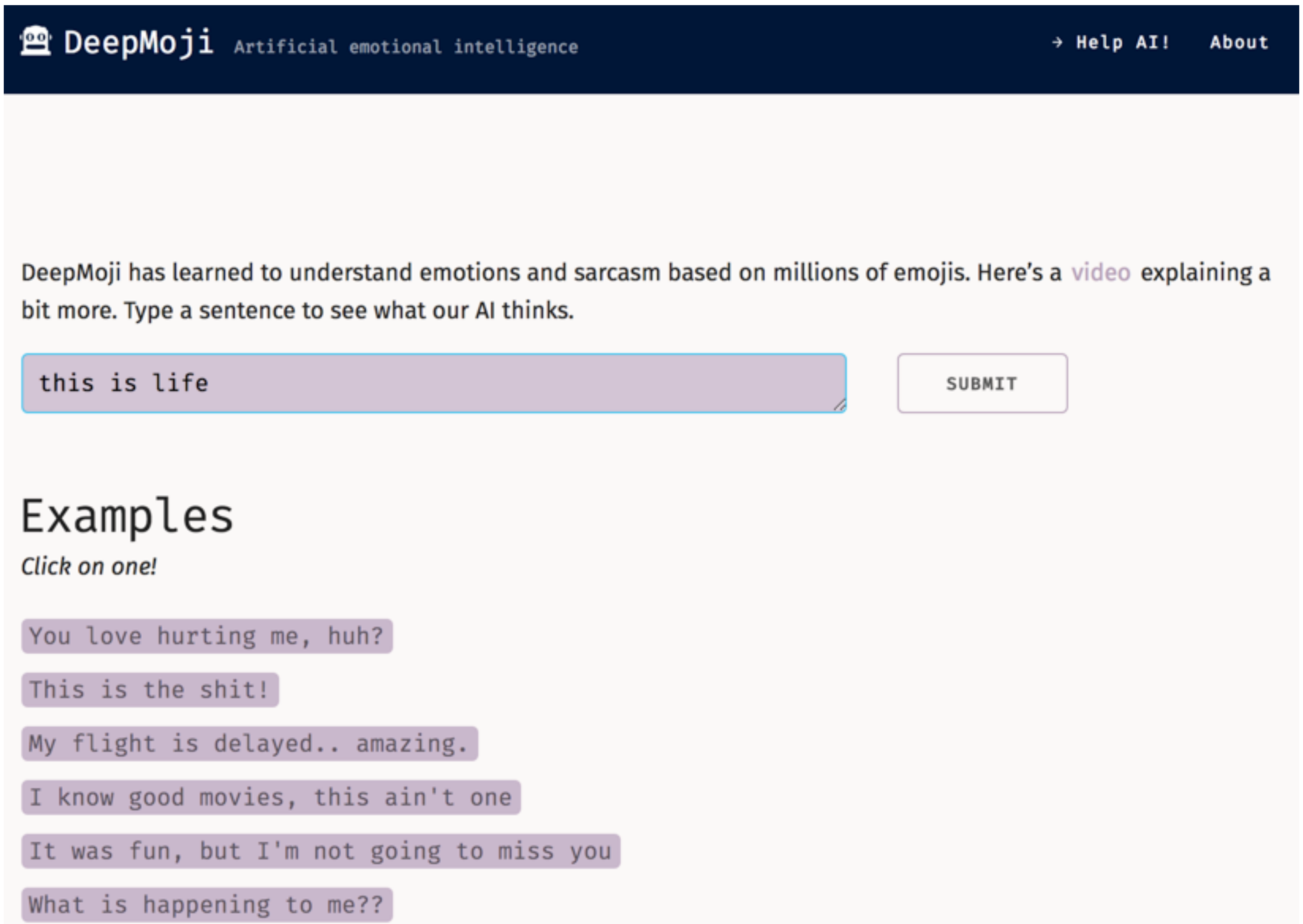
Still some fun

- Predict Emoji [24]



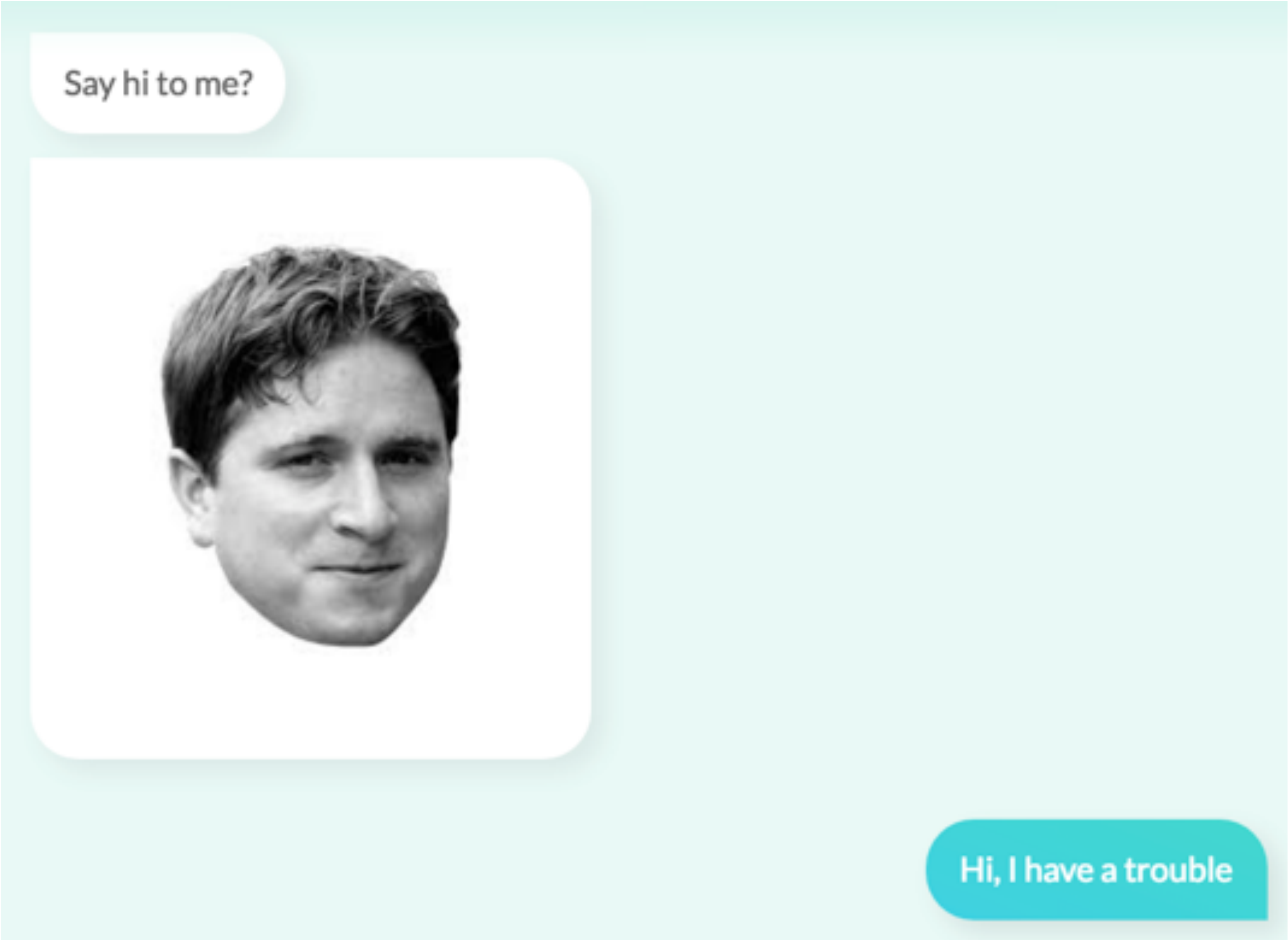
Still some fun

- Add this to your system [25]



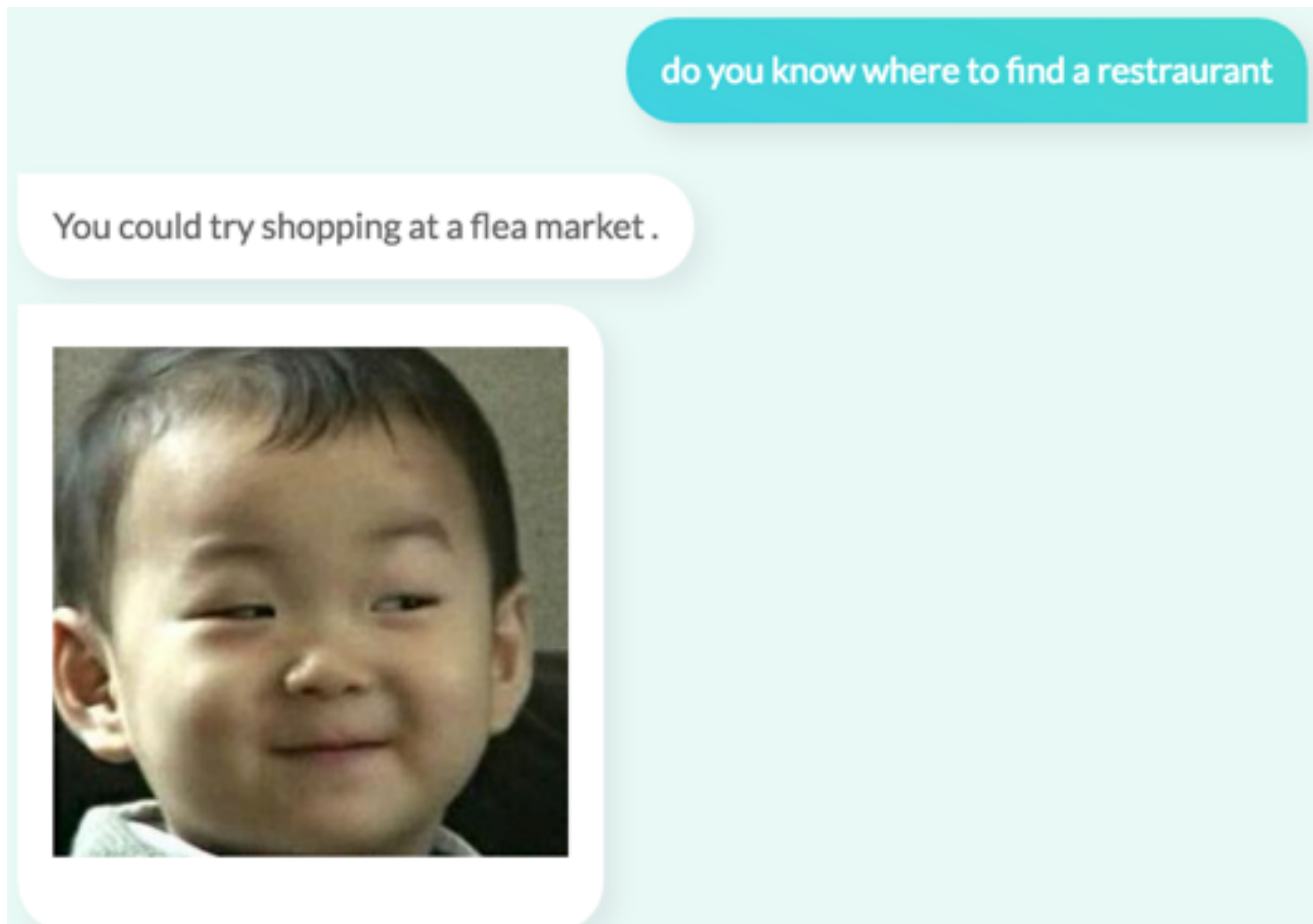
Still some fun

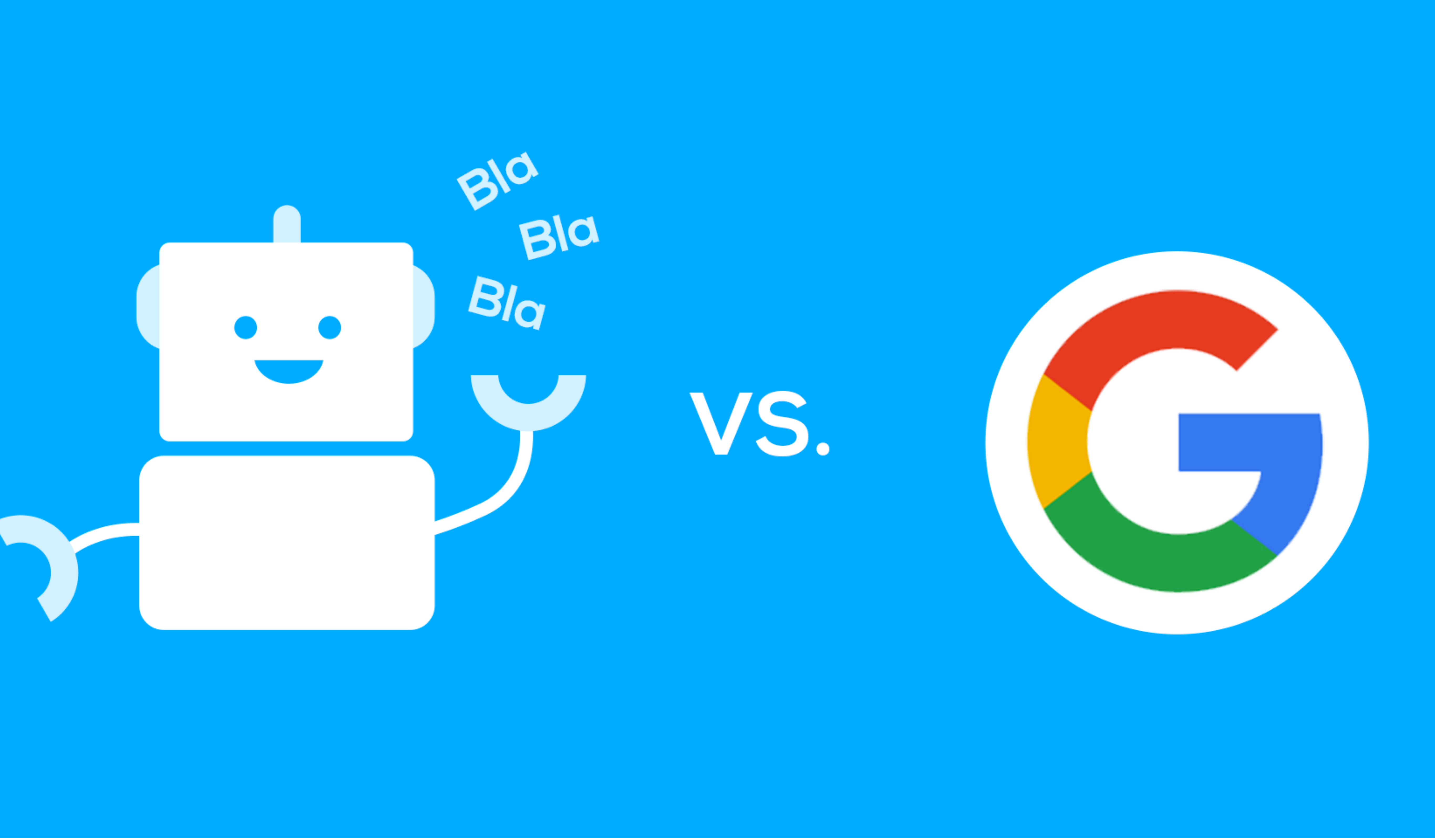
- Even cute features



Still some fun

- Even cute features





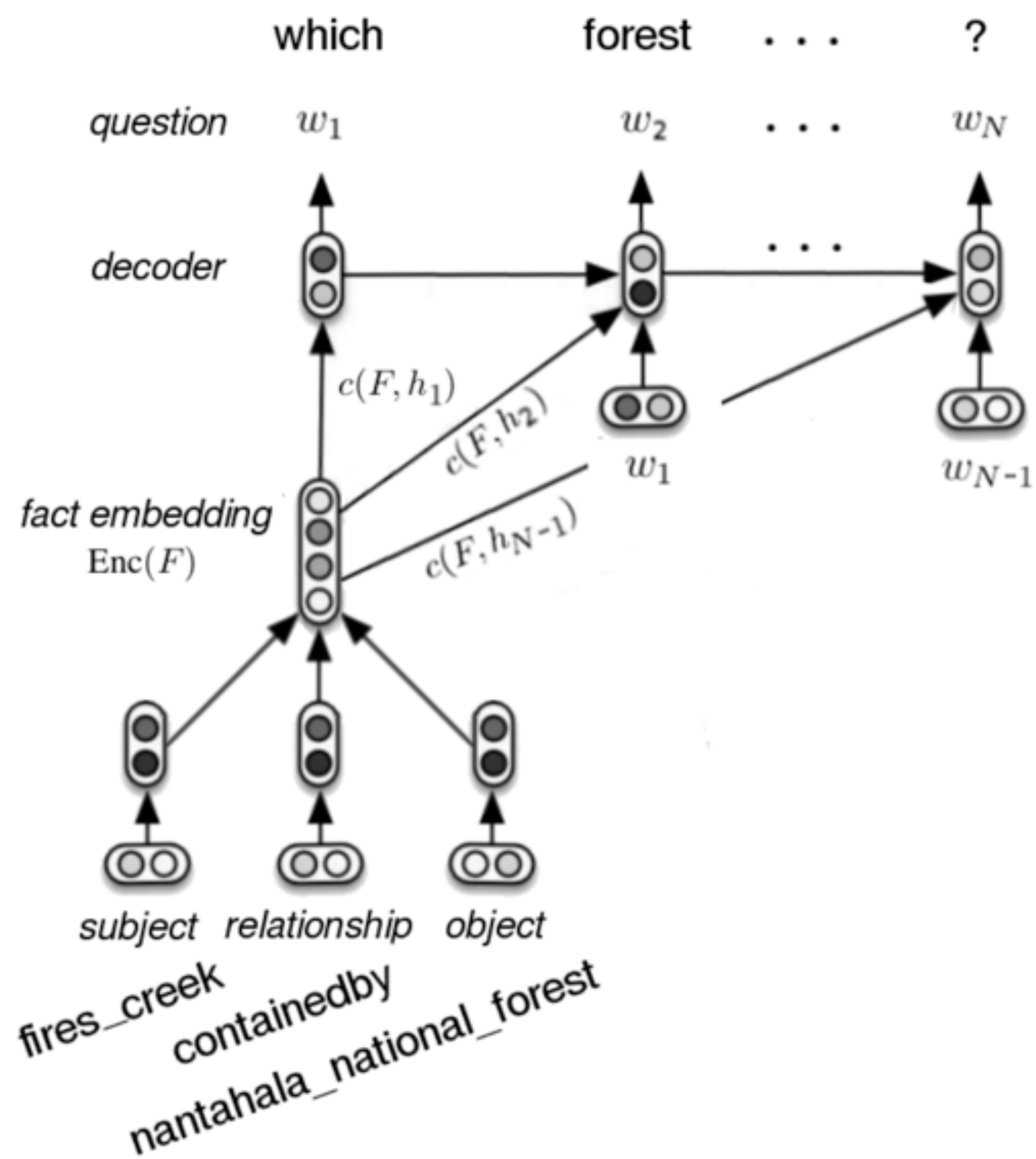
Knowledgable | Conversational Agents

Motivation

- Observation: The user inputs are often less informative [26]
- Inherent Incentives:
  - We often say short words and interact with others (communication theory)
  - We expect machines are equipped with (common) knowledge
  - We say words under some situations related to our intentions

Knowledge - Fact

- Uses attention-based RNN to generate factoid questions from Freebase triples [27]



Knowledge - Linguistics

- Linguistic Structure [28, 29]

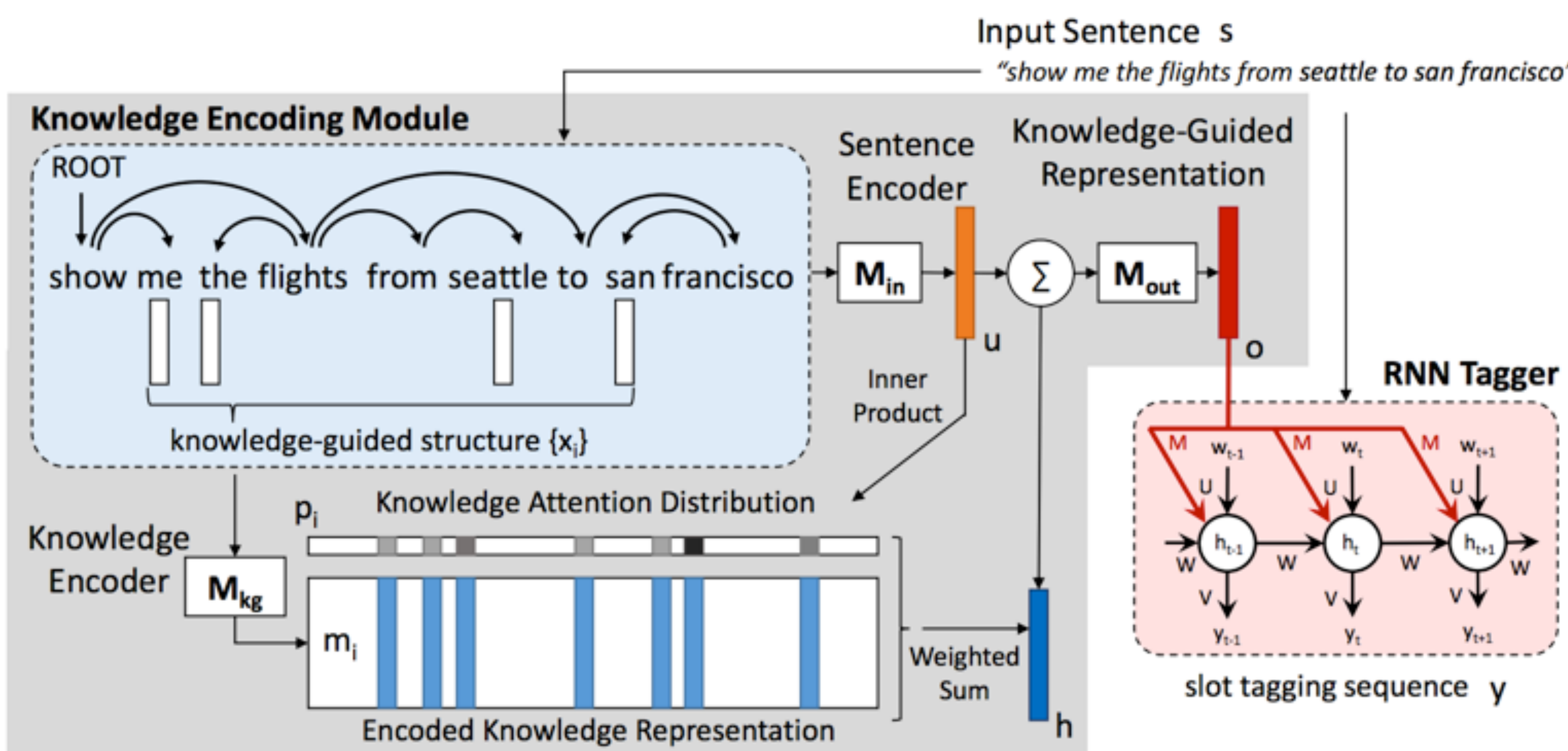
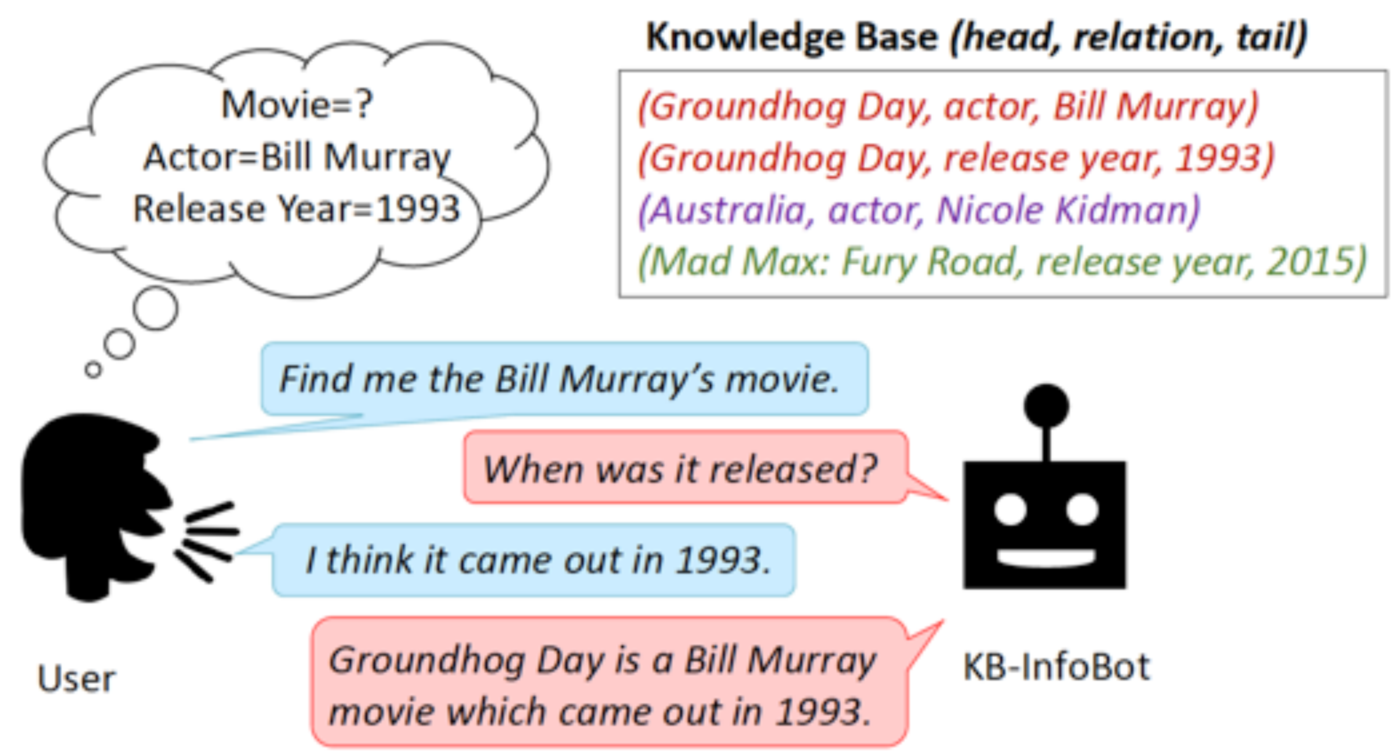


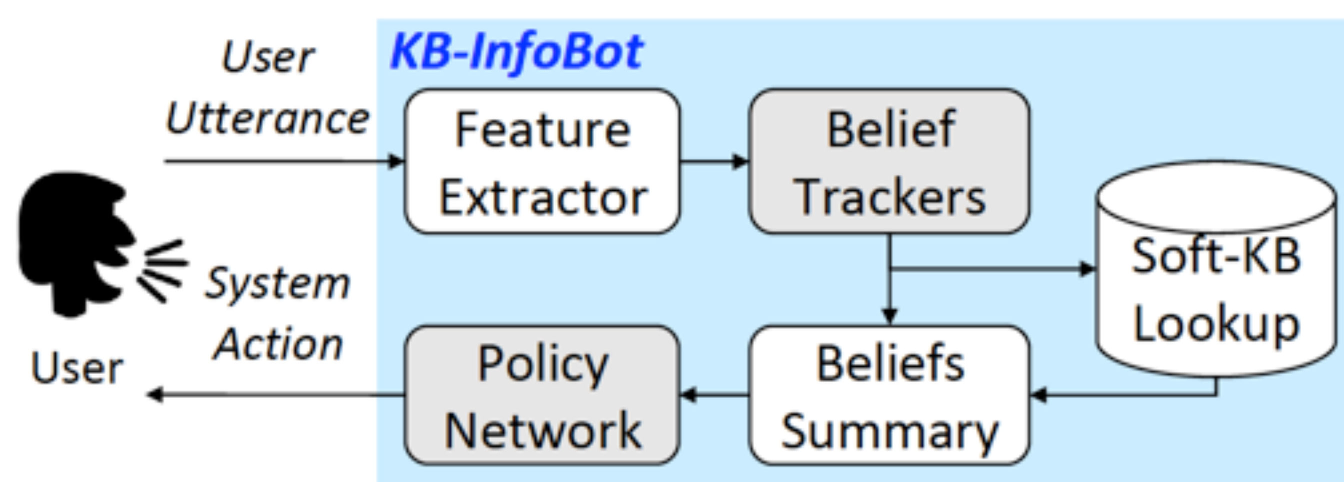
Figure 2: The illustration of knowledge-guided structural attention networks (K-SAN) for NLU.

Knowledge - Knowledge Base

- KB-InfoBot [30]

- a dialogue agent that provides users with an entity from a knowledge base (KB) by interactively asking for its attributes.



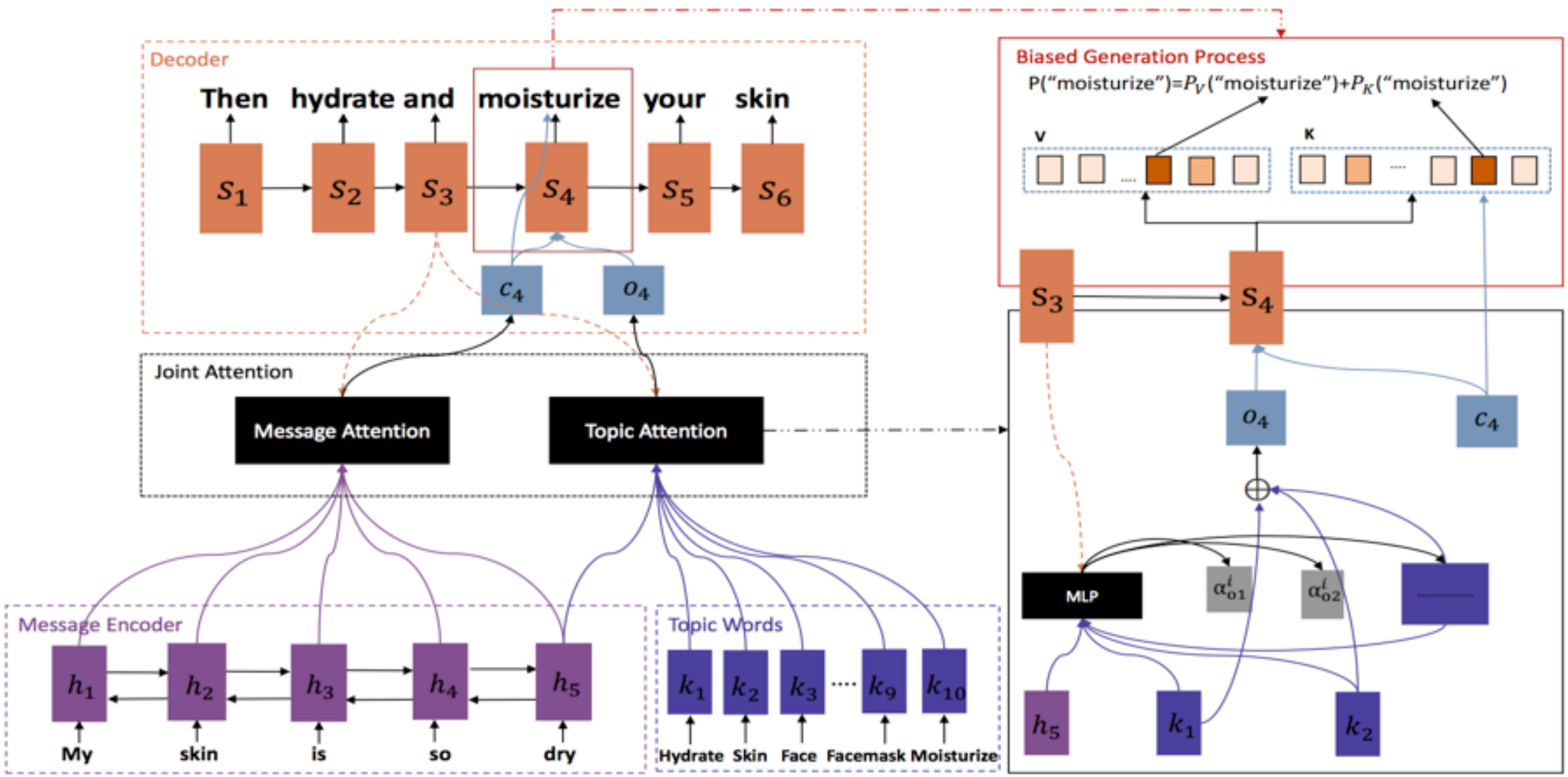


Knowledge - Topic

- utilize topics to simulate prior knowledge of human [31]
- leverage the topic information in generation by a joint attention mechanism and a biased generation probability.
- The joint attention mechanism summarises the hidden vectors of an input message as context vectors by message attention, synthesizes topic vectors by topic attention from the topic words of the message obtained from a pre-trained LDA model

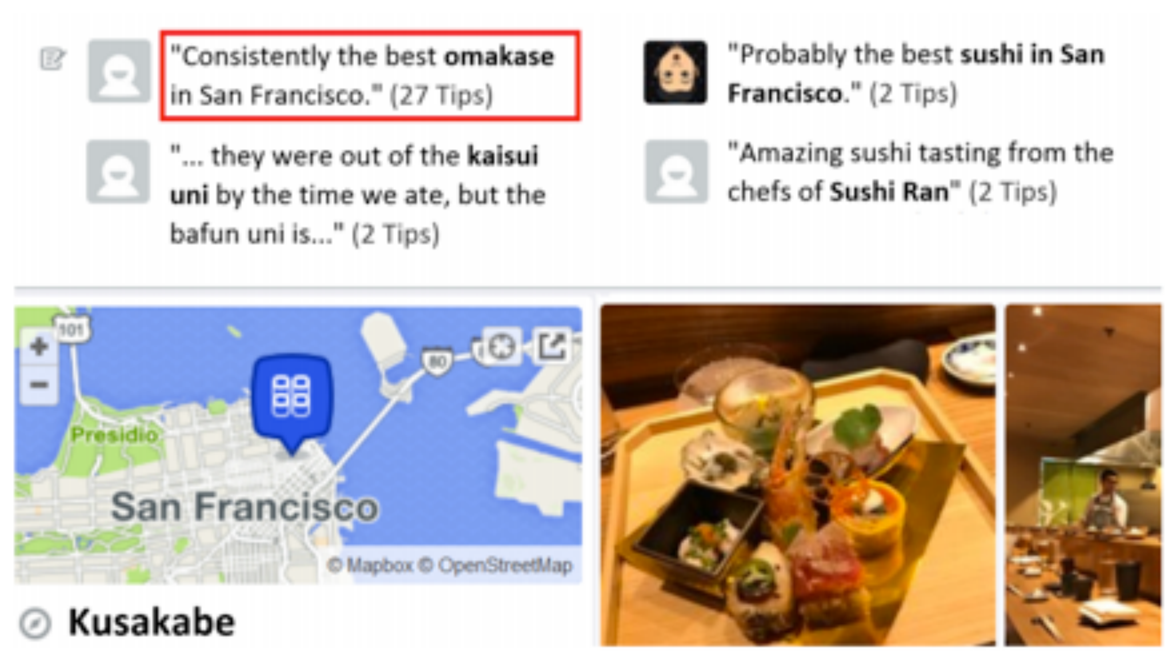
Knowledge - Topic

- utilize topics to simulate prior knowledge of human [31]



Knowledge - Scenario (Image, Articles…)

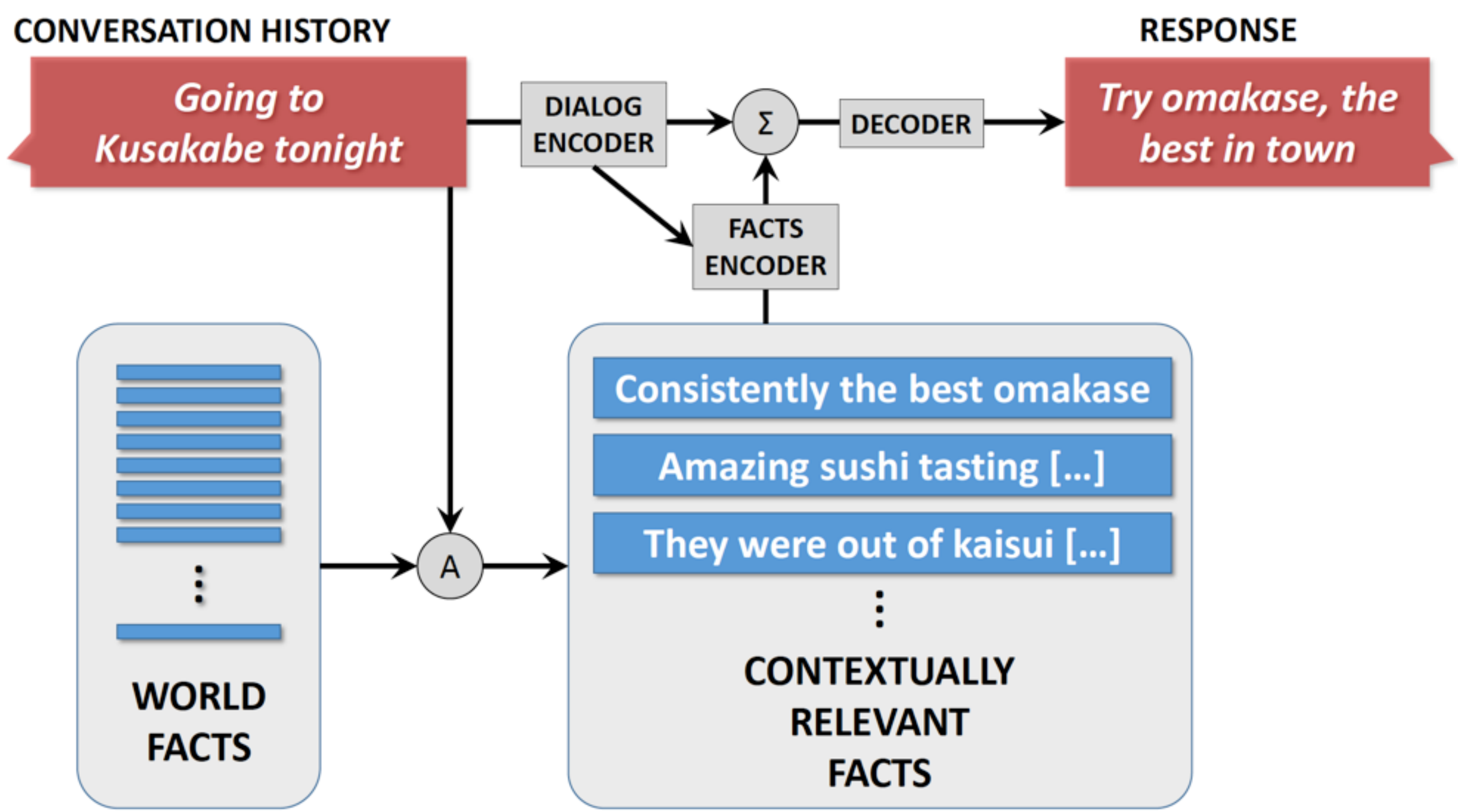
- incorporate factual information or entity-grounded opinion [32]
- producing more contentful responses without slot filling



User input: *Going to Kusakabe tonight.*
Neural model: Have a great time!
Human: You'll love it! Try omasake, the best in town.

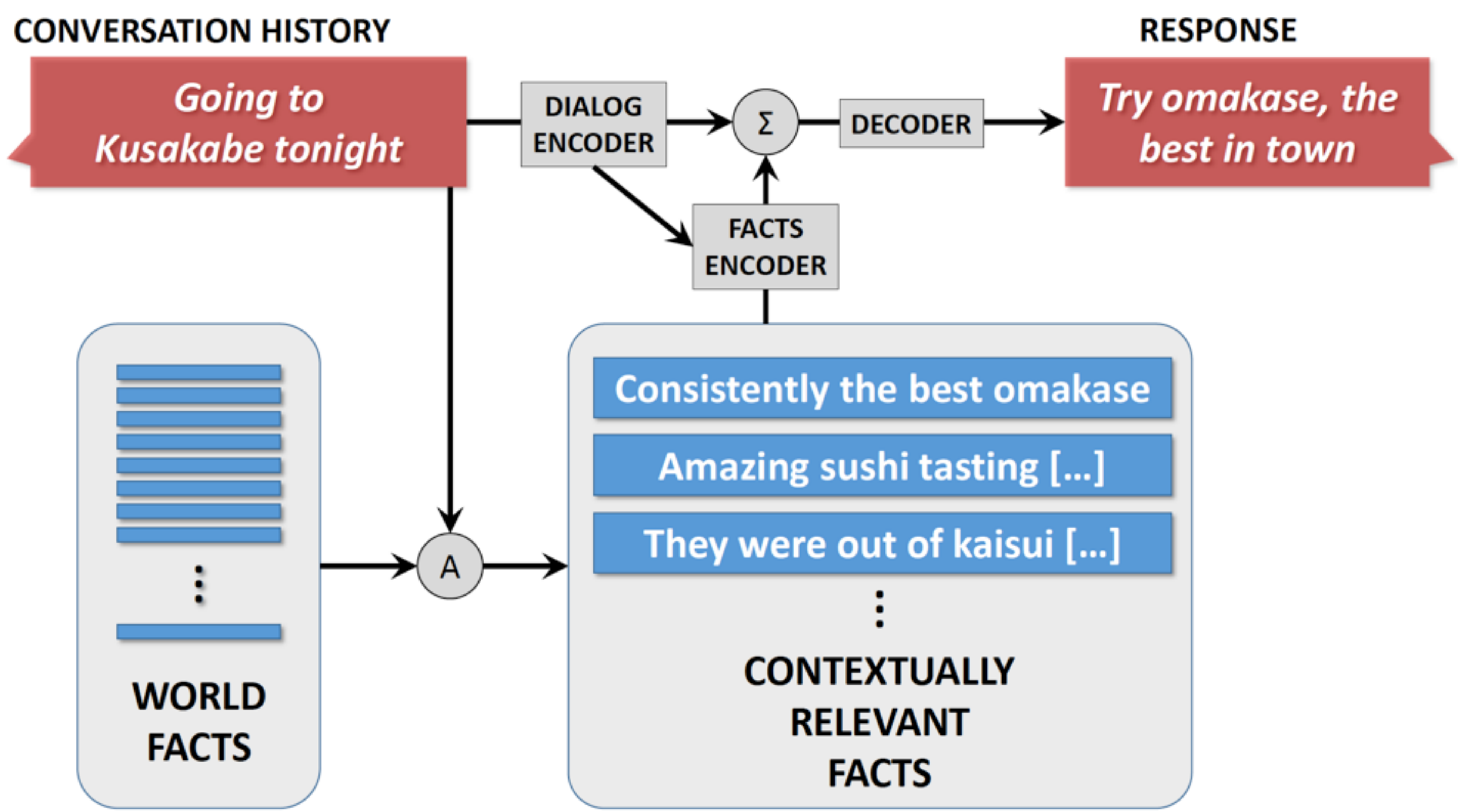
Knowledge - Scenario (Image, Articles…)

- incorporate factual information or entity-grounded opinion [32]



Knowledge - Scenario (Image, Articles…)

- incorporate factual information or entity-grounded opinion [32]

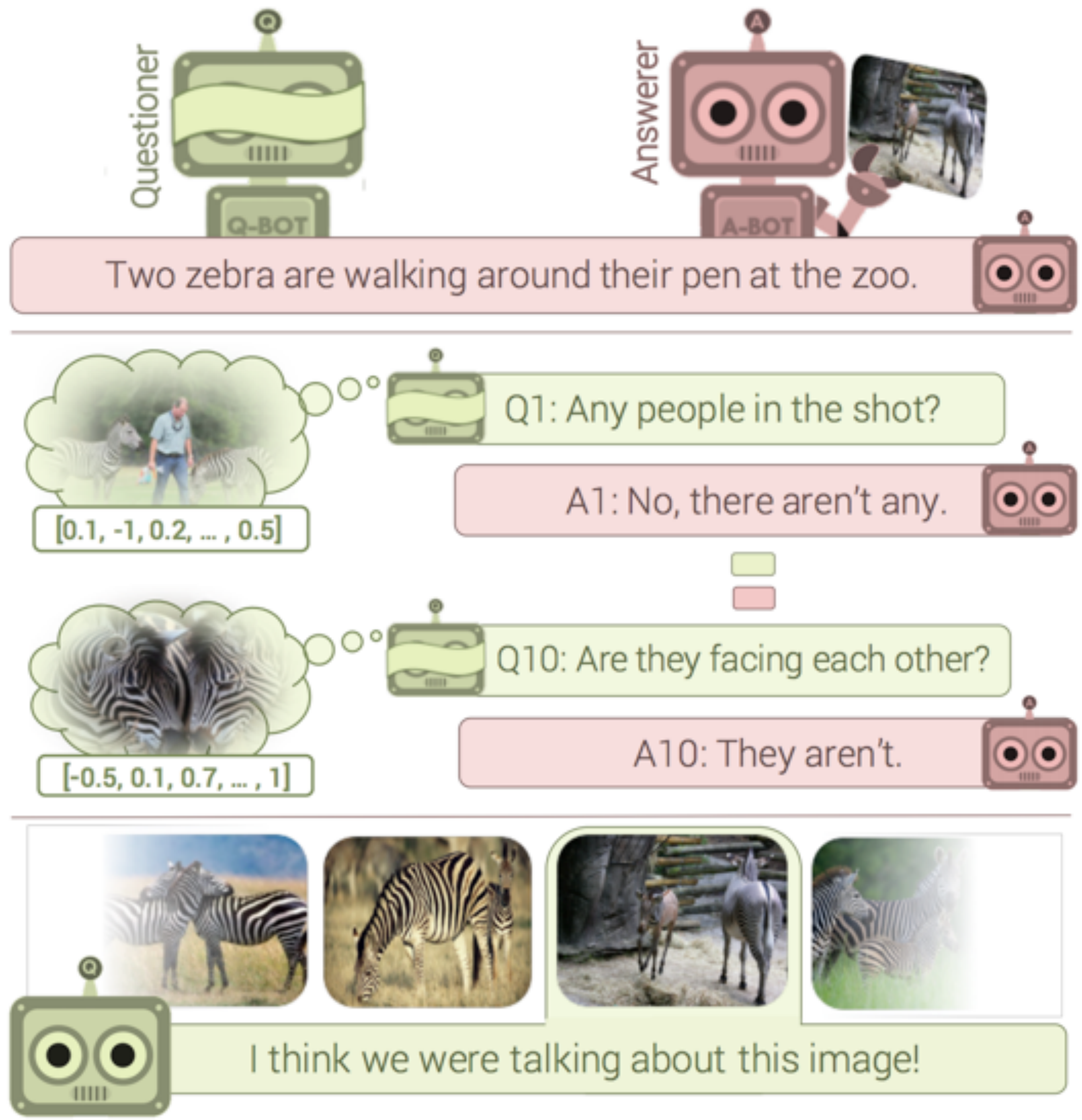


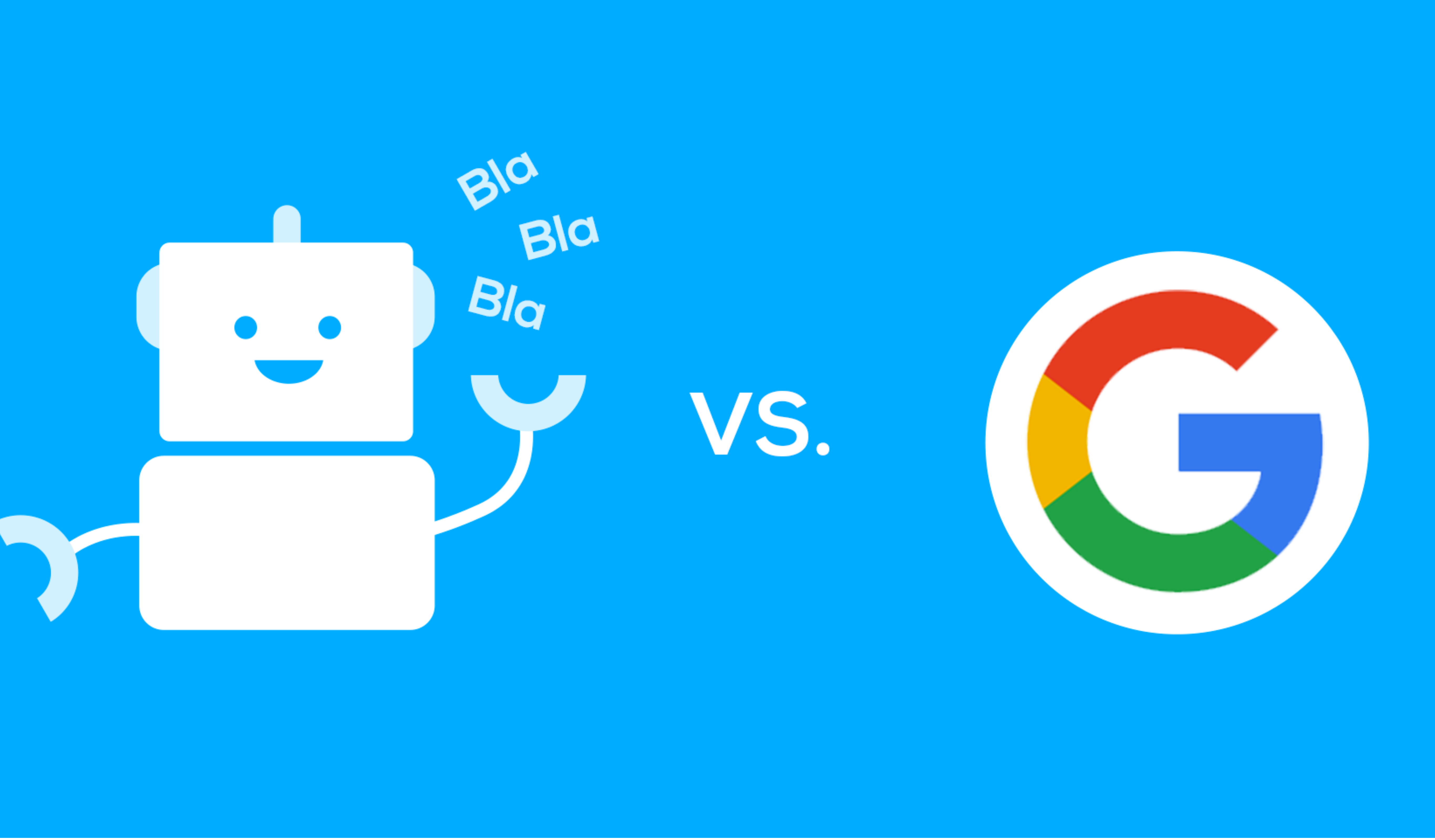
Knowledge - Scenario (Image, Articles…)

- Visual Dialog [32]

Knowledge - Scenario (Image, Articles…)

- Visual Dialog [33]




Bla
Bla
Bla
vs.

Dataset

- [1] full OpenSubtitle
- [2] partial OpenSubtitle
- [3] Sina Weibo dataset 2017, http://www.aihuang.org/p/challenge.html, http://tcci.ccf.org.cn/conference/2017/cfpt.php
- [4] Sina Weibo dataset 2016
- [5] Sina Weibo dataset 2015
- [6] Twitter data 2011
- [7] Chinese sentimental conversation

Dataset 2

- SQuAD
- Ubuntu Dialogue Corpus
- FAIR Movie Dialog Dataset
- Reddit Comment
- Microsoft Research Social Media Conversation Corpus
- Dialog State Tracking Challenge
- dgk_lost_conv 中文对白语料
- 高质量闲聊语料(需购买)
- Cornell Movie-Dialogs Corpus
- Maluuba NewsQA
- Maluuba Frames
- 百度 WebQA
- Noah NRM Data
- Noah Chinese QA

Try as a volunteer

- The Conversational Intelligence Challenge https://deeppavlov.github.io/convai/eval_howto/
  - Both human evaluators and dialogue agents complete the same task.
  - Connect randomly with a peer.
  - The peer might be a chat bot or other human user. No information about identity of the peer is provided.
  - Both parties are given a text of a recent news/wikipedia article.
  - Discuss content of the article with the peer as long as you wish.
  - Choose another news/wikipedia article and/or anonymous peer.

Homework

- Tutorial Reading:
  - Tutorial: Deep Reinforcement Learning. ICML 2016. http://icml.cc/2016/tutorials/deep_rl_tutorial.pdf
  - Introduction to Reinforcement Learning. Advanced Topics 2015 (COMPM050/COMPGI13) Reinforcement Learning (video available).
    - http://www0.cs.ucl.ac.uk/staff/d.silver/web/Teaching.html
    - http://www0.cs.ucl.ac.uk/staff/d.silver/web/Teaching_files/intro_RL.pdf
- Paper Reading:
  - Towards End-to-End Reinforcement Learning of Dialogue Agents for Information Access
  - Deal or No Deal? End-to-End Learning for Negotiation Dialogues
- Code Reading:
  - KB-InfoBot: https://github.com/MiuLab/KB-InfoBot

Thanks for your attention!

Q&A

References

- [1] Pascanu, Razvan, Tomas Mikolov, and Yoshua Bengio. "On the difficulty of training recurrent neural networks." International Conference on Machine Learning. 2013.
- [2] Mikolov, Tomas, et al. "Recurrent neural network based language model." Interspeech. Vol. 2. 2010.
- [3] Colah. “Understanding LSTMs”. http://colah.github.io/posts/2015-08-Understanding-LSTMs/
- [4] Sutskever, Ilya, Oriol Vinyals, and Quoc V. Le. "Sequence to sequence learning with neural networks." Advances in neural information processing systems. 2014.
- [5] Bahdanau, Dzmitry, Kyunghyun Cho, and Yoshua Bengio. "Neural machine translation by jointly learning to align and translate." arXiv preprint arXiv:1409.0473 (2014).
- [6] Vinyals, Oriol, and Quoc Le. "A neural conversational model." arXiv preprint arXiv:1506.05869 (2015).
- [7] Wen, Tsung-Hsien, et al. "A network-based end-to-end trainable task-oriented dialogue system." arXiv preprint arXiv:1604.04562 (2016).
- [8] Bowman, Samuel R., et al. "Generating sentences from a continuous space." arXiv preprint arXiv:1511.06349 (2015).
- [9] Hu, Baotian, et al. "Convolutional neural network architectures for matching natural language sentences." Advances in neural information processing systems. 2014.

References

- [10] Qiu, Xipeng, and Xuanjing Huang. "Convolutional Neural Tensor Network Architecture for Community-Based Question Answering." IJCAI. 2015.
- [11] Wu, Yu, et al. "Sequential Match Network: A New Architecture for Multi-turn Response Selection in Retrieval-based Chatbots." ACL 2017.
- [12] https://github.com/seatgeek/fuzzywuzzy
- [13] https://www.linkedin.com/pulse/duplicate-quora-question-abhishek-thakur
- [14] Is That a Duplicate Quora Question? https://www.linkedin.com/pulse/duplicate-quora-question-abhishek-thakur
- [15] https://github.com/HouJP/kaggle-quora-question-pairs/
- [16] https://pan.baidu.com/s/1dEV01gd
- [17] Lifeng Shang, Zhengdong Lu, Hang Li. “Neural Responding Machine for Short-Text Conversation”. EMNLP 2015.
- [18] Serban, Iulian Vlad, et al. "A Hierarchical Latent Variable Encoder-Decoder Model for Generating Dialogues." AAAI. 2017.
- [19] Serban, Iulian Vlad, et al. "Multiresolution Recurrent Neural Networks: An Application to Dialogue Response Generation." AAAI. 2017.
- [20] Alessandro Sordoni, Yoshua Bengio et al., “A Hierarchical Recurrent Encoder-Decoder for Context-Aware Generative Query Suggestion”. CIKM 2015 slides.

References

- [21] Zhou Hao et al, “Emotional Chatting Machine: Emotional Conversation Generation with Internal and External Memory”. arXiv preprint 2017.
- [22] Xiaoyu Shen, Hui Su, Yanran Li, Wenjie Li, Shuzi Niu, Yang Zhao, Akiko Aizawa, Guoping Long. “A Conditional Variational Framework for Dialog Generation”. ACL 2017.
- [23] Jessica Ficler, Yoav Goldberg. “Controlling Linguistic Style Aspects in Neural Language Generation”. arXiv preprint 2017.
- [24] Bjarke Felbo, Alan Mislove, Anders Søgaard, Iyad Rahwan, Sune Lehmann. “Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion and sarcasm”. EMNLP 2017.
- [25] DeepMoji. https://deepmoji.mit.edu/
- [26] Louis Shao, et al. “Generating Long and Diverse Responses with Neural Conversation Models”. arXiv preprint 2017.
- [27] Iulian Vlad Serban, et al. “Generating Factoid QuestionsWith Recurrent Neural Networks: The 30M Factoid Question-Answer Corpus”. ACL 2016a.
- [28] Yun-Nung Chen et al. “End-to-End Memory Networks with Knowledge Carryover for Multi-Turn Spoken Language Understanding”. arXiv preprint 2016b.
- [29] Yoon Kim et al. “Structured Attention Networks”. ICLR 2017.

References

- [30] Bhuwan Dhingra, Lihong Li, Xiujun Li, Jianfeng Gao, Yun-Nung Chen, Faisal Ahmed, Li Deng. “Towards End-to-End Reinforcement Learning of Dialogue Agents for Information Access”. ACL 2017.
- [31] Chen Xing et al. “Topic Aware Neural Response Generation”. arXiv preprint 2016.
- [32] Marjan Ghazvininejad et al. “A Knowledge-Grounded Neural Conversation Model”. arXiv preprint 2017.
- [33] Abhishek Das, Satwik Kottur, Khushi Gupta, Avi Singh, Deshraj Yadav, José M. F. Moura, Devi Parikh, Dhruv Batra. “Visual Dialog”. CVPR 2017.
- [34] Abhishek Das, Satwik Kottur, José M. F. Moura, Stefan Lee, Dhruv Batra. “Learning Cooperative Visual Dialog Agents with Deep Reinforcement Learning”. arXiv preprint 2017.