Learning from Dialogue after Deployment Feed Yourself, Chatbot!

Braden Hancock
Computer Science Dept.
Stanford University

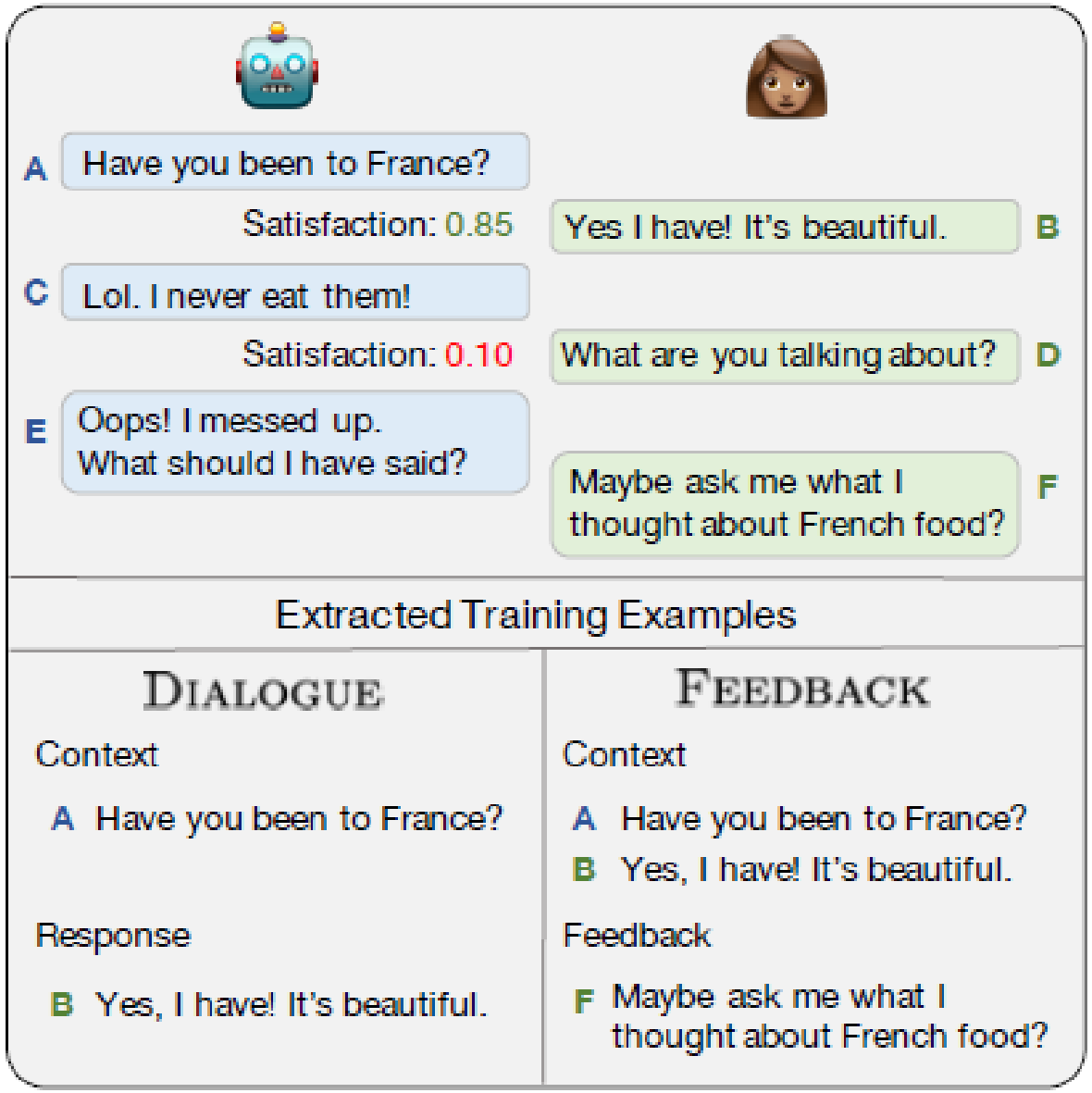
Antoine Bordes, Pierre-Emmanuel Mazare
Jason Weston
Facebook AI Research

Outline

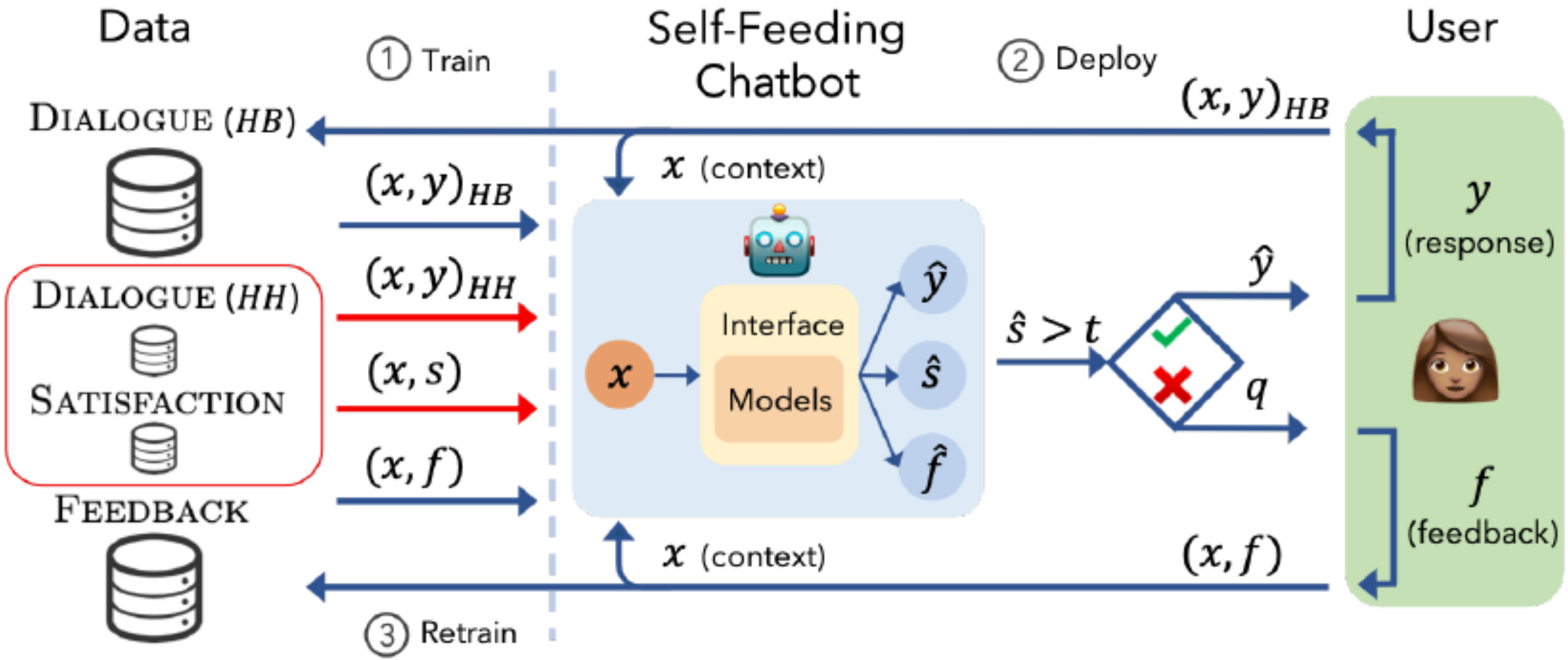
- Introduction
- Lifecycle of Self-Feeding Chatbot
  - Task 1: Dialogue
  - Task 2: Satisfaction
  - Task 3: Feedback
- Model Architecture
- Experimental Results
- Future Work

Introduction

- Fully supervised conversations tend to be expensive to collect in sufficient quantity;
- Naively training a dialogue agent on its own conversations yields poor results;
- Contributions:
  - proposed self-feeding chatbot;
  - dialogue ability improves by imitating human responses
  - classifying user satisfaction is important
  - release three new datasets


A
Have you been to France?
Satisfaction: 0.85
Yes I have! It's beautiful.
B
C
Lol. I never eat them!
Satisfaction: 0.10
What are you talking about?
D
E
Oops! I messed up.
What should I have said?
Maybe ask me what I thought about French food?
F
Extracted Training Examples
DIALOGUE
Context
A Have you been to France?
Response
B Yes, I have! It's beautiful.
FEEDBACK
Context
A Have you been to France?
B Yes, I have! It's beautiful.
Feedback
F Maybe ask me what I thought about French food?

Lifecycle of Self-Feeding Chatbot



Task 1: Dialogue

Task	Train	Valid	Test	Total
DIALOGUE				
– HH (HUMAN-HUMAN)	131438	7801	6634	145873
– HB (HUMAN-BOT)	60000	0	0	60000
FEEDBACK	60000	1000	1000	62000
SATISFACTION	1000	500	1000	2500

Task 2: Satisfaction

- Predict whether a speaking partner is satisfied with the quality of the current conversation;
- Examples take the form of (x, s) pairs;
- Dataset

Task 3: Feedback

- Predict the feedback that will be given by the speaking partner;
- Examples take the form of (x, f) pairs;
- Dataset

Examples of Types of Feedback

Category	%	Feedback Examples
Verbatim	53.0	• my favorite food is pizza • no, i have never been to kansas • i like when its bright and sunny outside
Suggestion	24.5	• you could say hey, i'm 30. how old are you? • yes, i play battlefield would have a been a great answer. • you could have said "yes, I'm happy it's friday."
Instructions	14.5	• tell me what your favorite breakfast food is • answer the question about having children! • tell me why your mom is baking bread
Options	8.0	• you could have said yes it really helps the environment or no its too costly • you could have said yes or no, or talked more about your mustang dream. • you should have said new york, texas or maryland. something like one of those.

Model Architecture

- Interface component
  - Shared by all tasks
- Model component
  - Contains a neural network for each task

Transformer

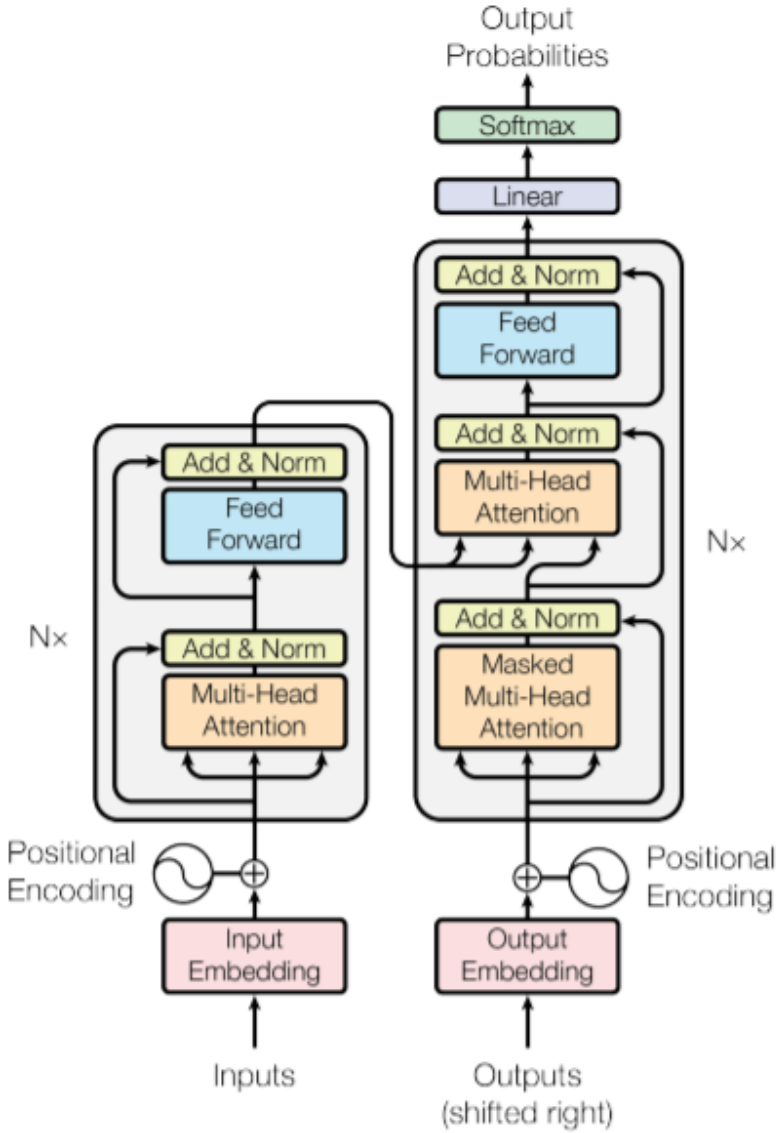


Figure 1: The Transformer - model architecture.

A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin. 2017. Attention is all you need. arXiv preprint arXiv:1706.03762.

Model Settings

- Maximum dialogue history length of 2;
- Tokens are embedded with fastText 300-dimensional embeddings;
- Vocabulary size: varies from 11.5k to 23.5k words;
- AdaMax optimizer;
- Batch size 128;
- Transformer layer: two attention heads and FFN size 32;
- Initial learning rate (0.001-0.005);
- Number of Transformer layers (1-2);
- Task-specific loss factors (0.5-2.0)

Experimental Results

Human-Bot (HB)		Human-Human (HH) Dialogue			
Dialogue	Feedback	20k	40k	60k	131k
-	-	30.3 (0.6)	36.2 (0.4)	39.1 (0.5)	44.7 (0.4)
20k	-	32.7 (0.5)	37.5 (0.6)	40.2 (0.5)	45.5 (0.7)
40k	-	34.5 (0.5)	37.8 (0.6)	40.6 (0.6)	45.1 (0.6)
60k	-	35.4 (0.4)	37.9 (0.7)	40.2 (0.8)	45.0 (0.7)
-	20k	35.0 (0.5)	38.9 (0.3)	41.1 (0.5)	45.4 (0.8)
-	40k	36.7 (0.7)	39.4 (0.5)	41.8 (0.4)	45.7 (0.6)
-	60k	37.8 (0.6)	40.6 (0.5)	42.2 (0.7)	45.8 (0.7)
60k	60k	**39.7** (0.6)	**42.0** (0.6)	**43.3** (0.7)	**46.3** (0.8)

F1 Score on Satisfaction Task

Method	Pr.	Re.	F1
Uncertainty Top	0.39	0.99	0.56
(Pr. $\geq$ 0.5)	0.50	0.04	0.07
Uncertainty Gap	0.38	1.00	0.55
(Pr. $\geq$ 0.5)	0.50	0.04	0.07
Satisfaction Regex	0.91	0.27	0.42
Satisfaction Classifier (1k)	0.84	0.84	0.84
Satisfaction Classifier (2k)	0.89	0.84	0.87
Satisfaction Classifier (5k)	0.94	0.82	0.88
Satisfaction Classifier (20k)	0.96	0.84	0.89
Satisfaction Classifier (40k)	0.96	0.84	**0.90**

Future Work

- More ways a model could learn to improve itself:
  - E.g. learning which question to ask in a given context to receive the most valuable feedback
- Use the flexible nature of dialogue to intermix data collection of more than one type
  - sometimes requesting new FEEDBACK examples;
  - other times requesting new SATISFACTION examples (e.g., asking “Did my last response make sense?”

References

- Learning from Dialogue after Deployment Feed Yourself, Chatbot!
- Paper: https://arxiv.org/pdf/1901.05415.pdf
- Code: https://parl.ai/projects/self_feeding/
- A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin. 2017. Attention is all you need. arXiv preprint arXiv:1706.03762.
- Link: https://arxiv.org/pdf/1706.03762.pdf