Deep Active Learning for Dialogue Generation

Nabiha Asghar

6th Joint Conference on Lexical and Computational Semantics (*SEM 2017)

Joint work with:
Pascal Poupart, Xin Jiang (Huawei) and Hang Li (Huawei)

2016
The Year of Chatbots for Businesses

JAN

WhatsApp launches plans to allow businesses on their network

MAR

Microsoft launches **Bot Framework**

APR

Kik launches **Bot Shop**

Facebook launches **Messenger Platform** for bots

SEP

Oracle announces new chatbot platform

Google announces acquisition of chatbot maker **API.ai**

Google launches **Google Assistant** inside new **Allo app**

Amazon acqui-hires chatbot maker **Angel.ai**

LinkedIn announces introduction of first messaging bot

LINE launches new messaging API

NOV

Viber launches Public Accounts to chat with businesses

Amazon launches chatbot platform **Amazon Lex**

DEC

Google launches API for **Google Home**

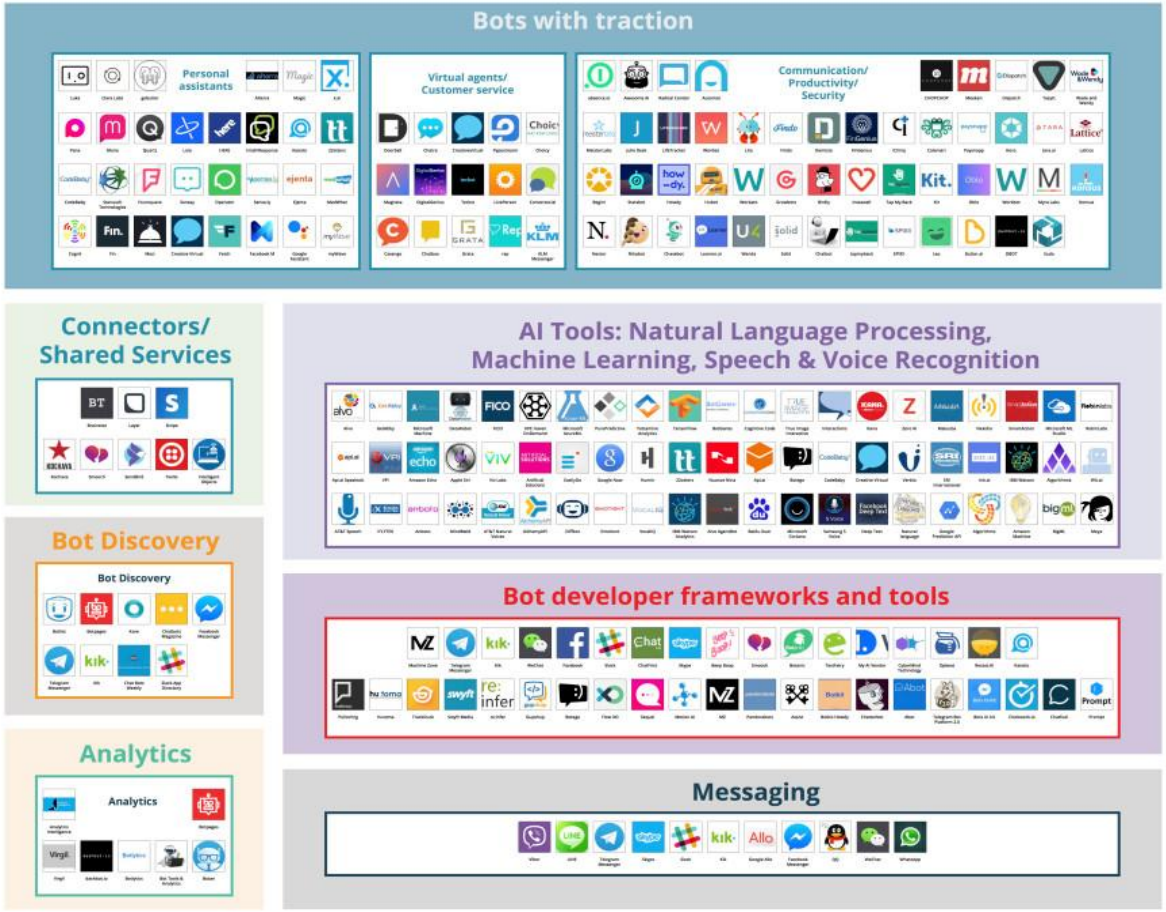
aspect.

Source: http://blogs.aspect.com/2016-the-year-of-the-chatbot/

DESIGNED BY
JON CIFUENTES

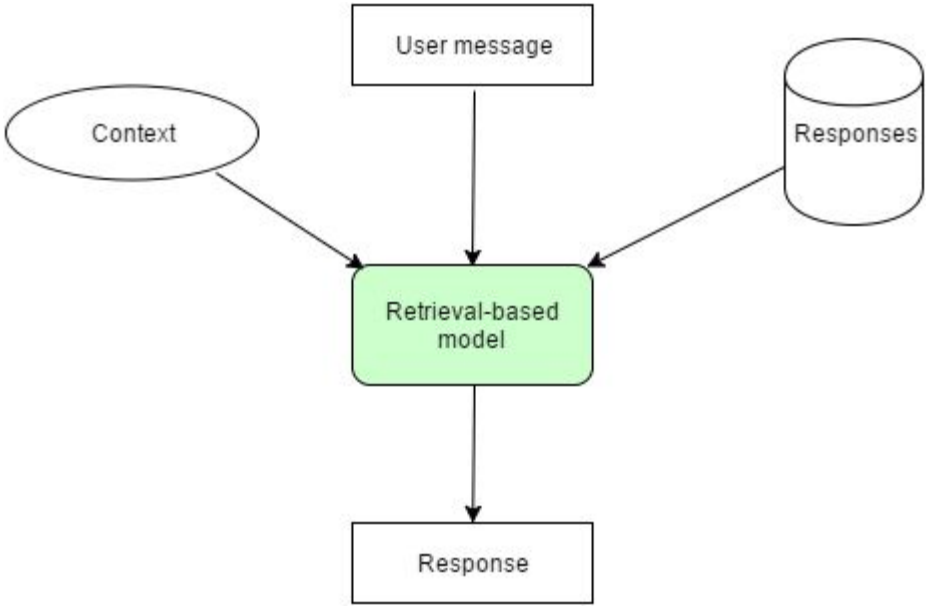
Bots Landscape

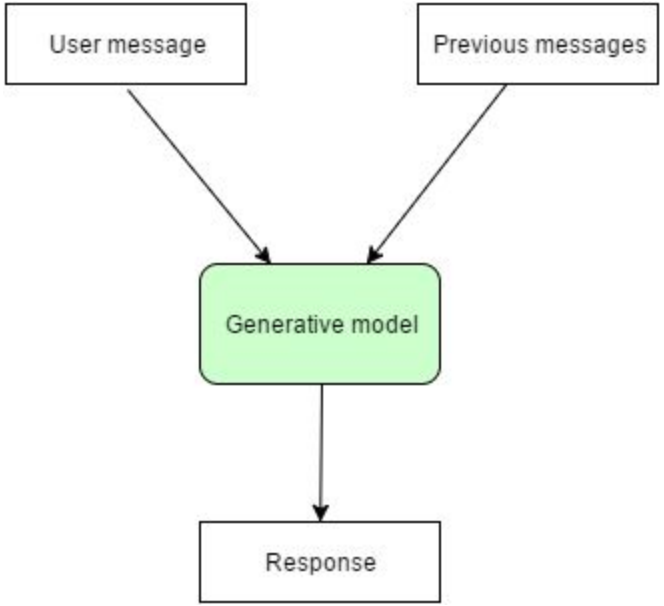
POWERED BY
VB | Profiles



Source:
https://venturebeat.com/2016/08/11/introducing-the-bots-landscape-170-companies-4-billion-in-funding-thousands-of-bots/

Retrieval-based vs. Generative





Source: https://hackernoon.com/chatbot-architecture-496f5bf820ed

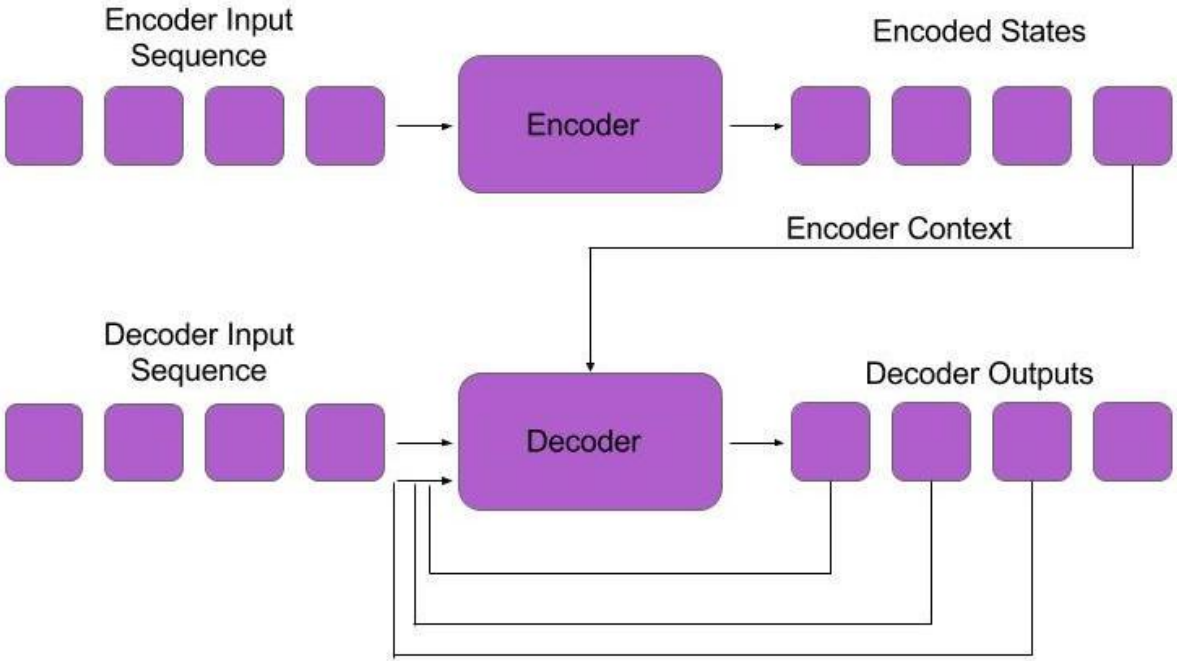
Premature Rise of Chatbots

Facebook Inc's Chatbots Hit a 70% Failure Rate (The Motley Fool)

The problem with chatbots: How to make them more human? (Chatbotsmagazine.com)

Bots are stumbling (McKinsey&Company)

Generative Models: Seq2Seq (Google)



Source: https://indico.io/blog/sequence-modeling-neuralnets-part1/

Generative Models (Seq2Seq)

me: Hello?
bot: Hi.
• me: How are you?
bot: I'm fine.
• me: What's your name?
bot: It's hard to describe.
•me: How so?
bot: I'm not sure.
• me: Are you intelligent?
bot: Yes, well...
• me: What are you?
bot: I'm not sure.
• me: Do you plan on taking over the world?
bot: No, i don't.

- Short, dull & generic responses
- Cannot keep the users engaged
- Infinite loops of repetition
- Mutually inconsistent responses

Idea: Reinforcement/Active Learning

Goal: Incorporate user feedback into the conversation: Human In The Loop

Research question: how do we construct a comprehensive and effective feedback mechanism for generative dialogue models?

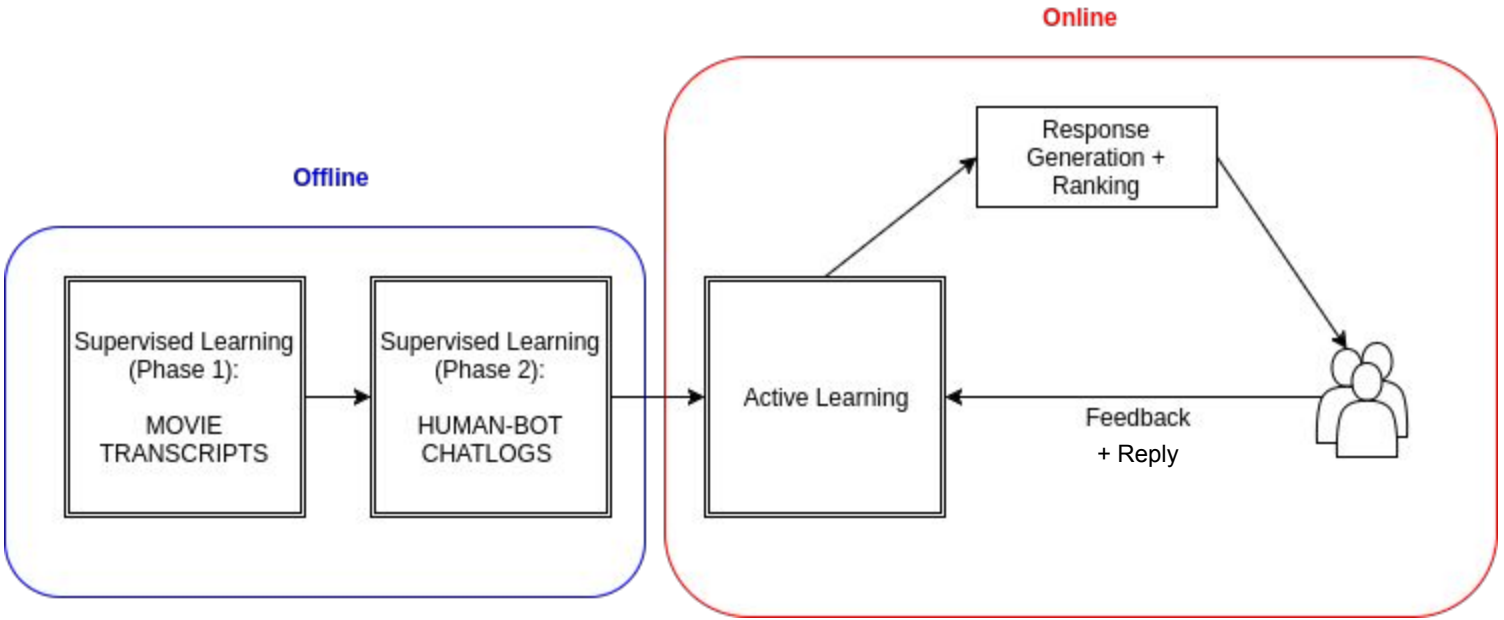
Existing DRL Models for CAs

- Task-oriented CAs: reward function is often obvious
- Open-domain CAs: hand-crafted reward functions (ease of answering, informativity, diversity, etc)

Existing DRL Models for CAs

- Task-oriented CAs: reward function is often obvious
- Open-domain CAs: hand-crafted reward functions (ease of answering, informativity, diversity, etc)
- Our work:
    - Builds on top of a trained Seq2Seq model
    - A simple, all-encompassing, user-centric feedback mechanism
    - Online, incremental learning at each turn in the conversation
    - No need to hand-craft the reward function

Training Pipeline



Example Transcript

Human user: I don't want to work today.

Example Transcript

Human user: I don't want to work today.
bot: Ok.
bot: I wish i could robot.
bot: Why not?
bot: Aww:).
bot: No, you've just got to be quick.

Example Transcript

Human user: I don't want to work today.
bot: Ok.
bot: I wish i could robot.
bot: Why not?
bot: Aww:).
bot: No, you've just got to be quick.
human feedback: 3

Example Transcript

Human user: I don't want to work today.
bot: Ok.
bot: I wish i could robot.
bot: Why not?
bot: Aww:).
bot: No, you've just got to be quick.
human feedback: 3
Human user: I just want to lie in my bed all day and have pizza.

Example Transcript

Human user: I don't want to work today.
bot: Ok.
bot: I wish i could robot.
bot: Why not?
bot: Aww:).
bot: No, you've just got to be quick.
human feedback: 3
Human user: I just want to lie in my bed all day and have pizza.
bot: I can't do that again!
bot: Does he _ _ _.
bot: Good idea.
bot: Ok.
bot: Why not?

Example Transcript

Human user: I don't want to work today.
bot: Ok.
bot: I wish i could robot.
bot: Why not?
bot: Aww:).
bot: No, you've just got to be quick.
human feedback: 3
Human user: I just want to lie in my bed all day and have pizza.
bot: I can't do that again!
bot: Does he _ _ _.
bot: Good idea.
bot: Ok.
bot: Why not?
human feedback: Good idea. May I join you?

Example Transcript

Human user: I don't want to work today.
bot: Ok.
bot: I wish i could robot.
bot: Why not?
bot: Aww:).
bot: No, you've just got to be quick.
human feedback: 3
Human user: I just want to lie in my bed all day and have pizza.
bot: I can't do that again!
bot: Does he _ _ _.
bot: Good idea.
bot: Ok.
bot: Why not?
human feedback: Good idea. May I join you?

Human user: I don't want to work today.

Example Transcript

Human user: I don't want to work today.
bot: Ok.
bot: I wish i could robot.
bot: Why not?
bot: Aww:).
bot: No, you've just got to be quick.
human feedback: 3
Human user: I just want to lie in my bed all day and have pizza.
bot: I can't do that again!
bot: Does he _ _ _.
bot: Good idea.
bot: Ok.
bot: Why not?
human feedback: Good idea. May I join you?

Human user: I don't want to work today.
bot: Why not?
bot: Aww:).
bot: Ok?
bot: I wish i could robot.
bot: No, you've just got to be quick.

Example Transcript

Human user: I don't want to work today.
bot: Ok.
bot: I wish i could robot.
bot: Why not?
bot: Aww:).
bot: No, you've just got to be quick.
human feedback: 3
Human user: I just want to lie in my bed all day and have pizza.
bot: I can't do that again!
bot: Does he _ _ _.
bot: Good idea.
bot: Ok.
bot: Why not?
human feedback: Good idea. May I join you?

Human user: I don't want to work today.
bot: Why not?
bot: Aww:).
bot: Ok?
bot: I wish i could robot.
bot: No, you've just got to be quick.
human feedback:
Human user: I just want to lie in my bed all day and have pizza.

Example Transcript

Human user: I don't want to work today.
bot: Ok.
bot: I wish i could robot.
bot: Why not?
bot: Aww:).
bot: No, you've just got to be quick.
human feedback: 3
Human user: I just want to lie in my bed all day and have pizza.
bot: I can't do that again!
bot: Does he _ _ _.
bot: Good idea.
bot: Ok.
bot: Why not?
human feedback: Good idea. May I join you?

Human user: I don't want to work today.
bot: Why not?
bot: Aww:).
bot: Ok?
bot: I wish i could robot.
bot: No, you've just got to be quick.
human feedback:
Human user: I just want to lie in my bed all day and have pizza.
bot: Good idea. May I join you?
bot: Does it has been to the?
bot: I can’t do that again!
bot: So good.
bot: Where good?

Example Transcript

Human user: I don't want to work today.
bot: Ok.
bot: I wish i could robot.
bot: Why not?
bot: Aww:).
bot: No, you've just got to be quick.
human feedback: 3
Human user: I just want to lie in my bed all day and have pizza.
bot: I can't do that again!
bot: Does he _ _ _.
bot: Good idea.
bot: Ok.
bot: Why not?
human feedback: Good idea. May I join you?

Human user: I don't want to work today.
bot: Why not?
bot: Aww:).
bot: Ok?
bot: I wish i could robot.
bot: No, you've just got to be quick.
human feedback:
Human user: I just want to lie in my bed all day and have pizza.
bot: Good idea. May I join you?
bot: Does it has been to the?
bot: I can’t do that again!
bot: So good.
bot: Where good?
human feedback:
Human user: Haha! Which pizza toppings are your favourite?

Active Learning at each step

For the message-response pair *(x,y)*, use cross-entropy loss to maximize the likelihood of *y* given *x*.

Active Learning at each step

For the message-response pair *(x,y)*, use cross-entropy loss to maximize the likelihood of *y* given *x*.

Use optimizer's *learning rate* to determine how quickly you want the bot to learn.

Active Learning at each step

For the message-response pair *(x,y)*, use cross-entropy loss to maximize the likelihood of *y* given *x*.

Use optimizer's *learning rate* to determine how quickly you want the bot to learn.

Feedback is optional (i.e. model/policy does not change).

Active Learning at each step

For the message-response pair *(x,y)*, use cross-entropy loss to maximize the likelihood of *y* given *x*.

Use optimizer's *learning rate* to determine how quickly you want the bot to learn.

Feedback is optional (i.e. model/policy does not change).

Can learn from multiple users simultaneously.

Active Learning at each step

For the message-response pair *(x,y)*, use cross-entropy loss to maximize the likelihood of *y* given *x*.

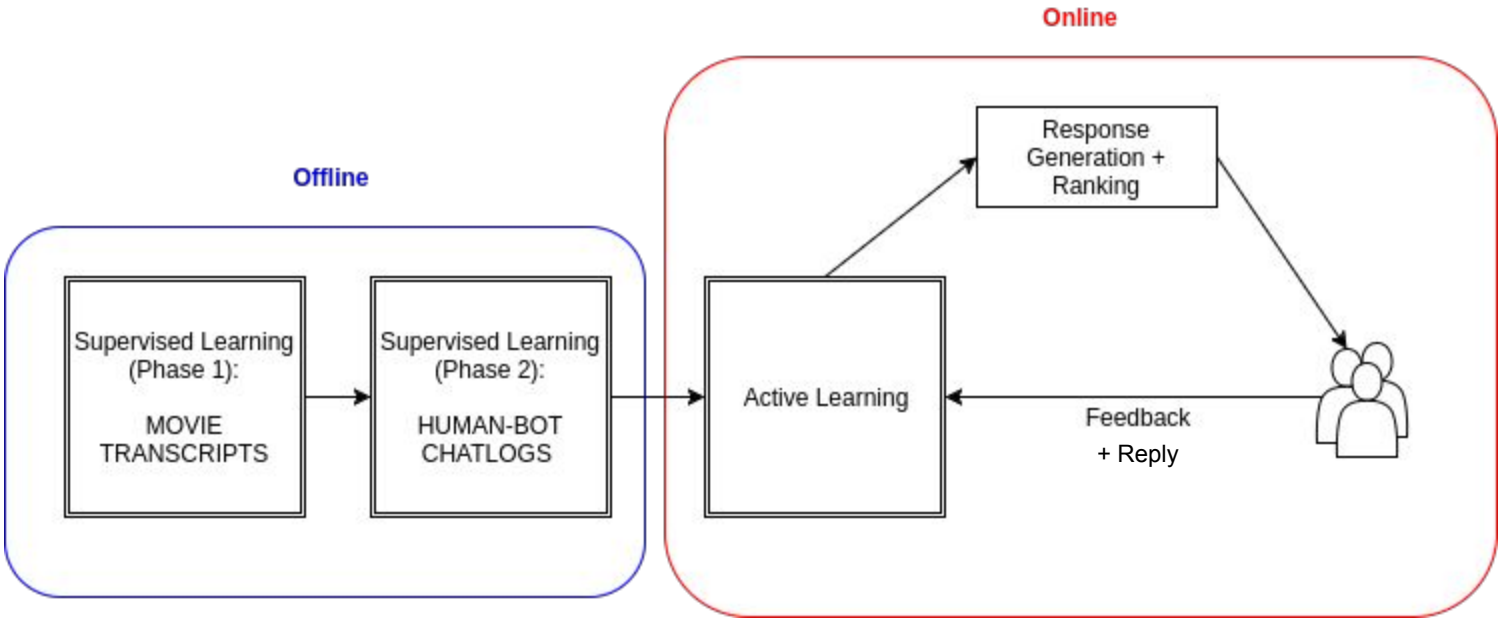
Use optimizer's *learning rate* to determine how quickly you want the bot to learn.

Feedback is optional (i.e. model/policy does not change).

Can learn from multiple users simultaneously.

Goal: generate interesting, relevant and diverse responses.

Training Pipeline



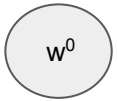
Heuristic Response Generation

Heuristic Response Generation

BEAM SEARCH

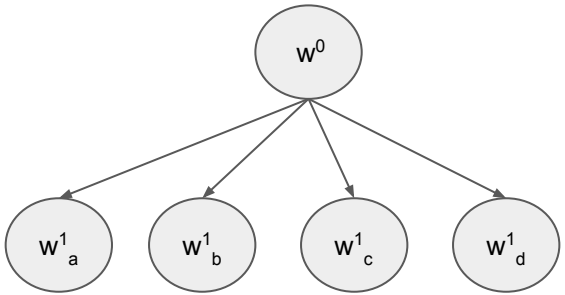
Heuristic Response Generation

BEAM SEARCH



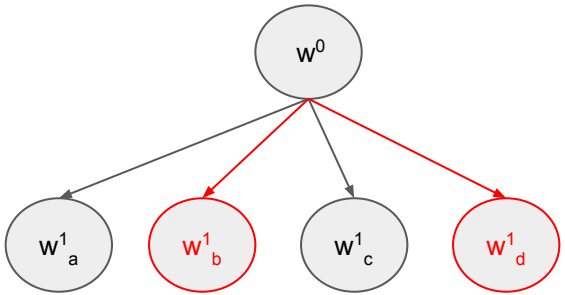
Heuristic Response Generation

BEAM SEARCH



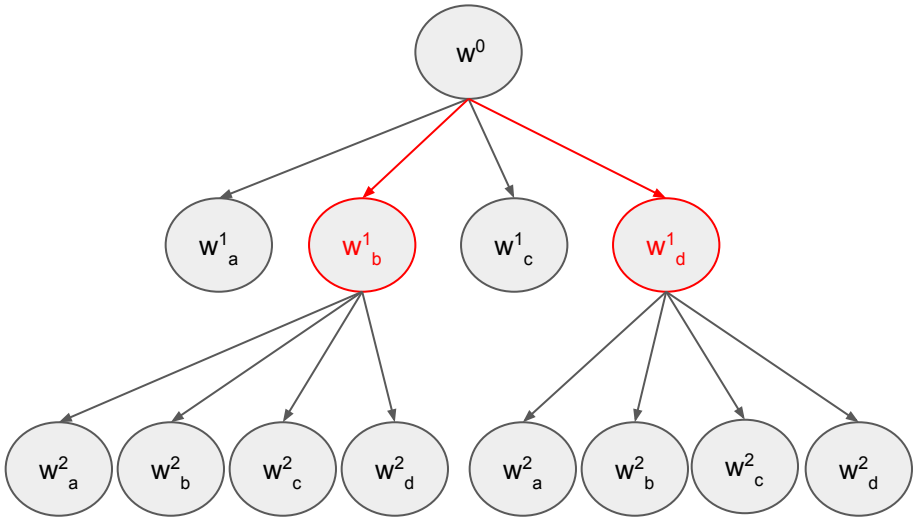
Heuristic Response Generation

BEAM SEARCH



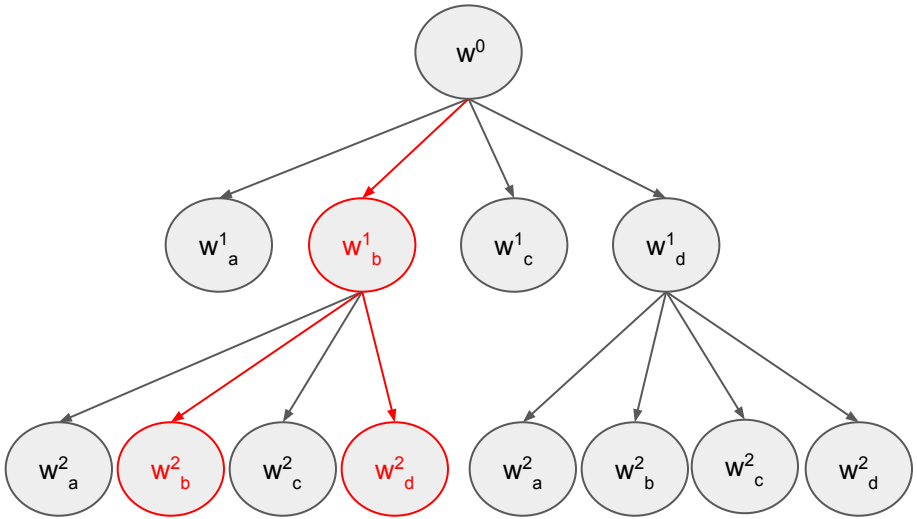
Heuristic Response Generation

BEAM SEARCH



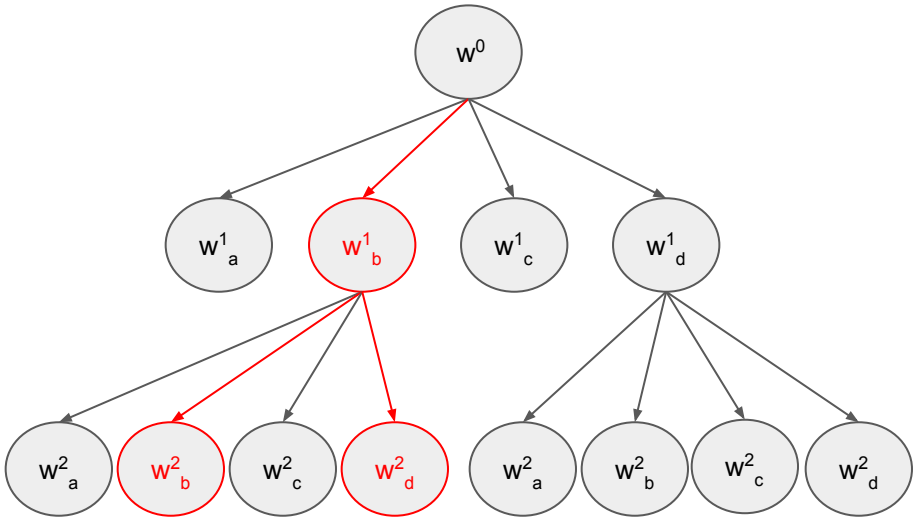
Heuristic Response Generation

BEAM SEARCH



Heuristic Response Generation

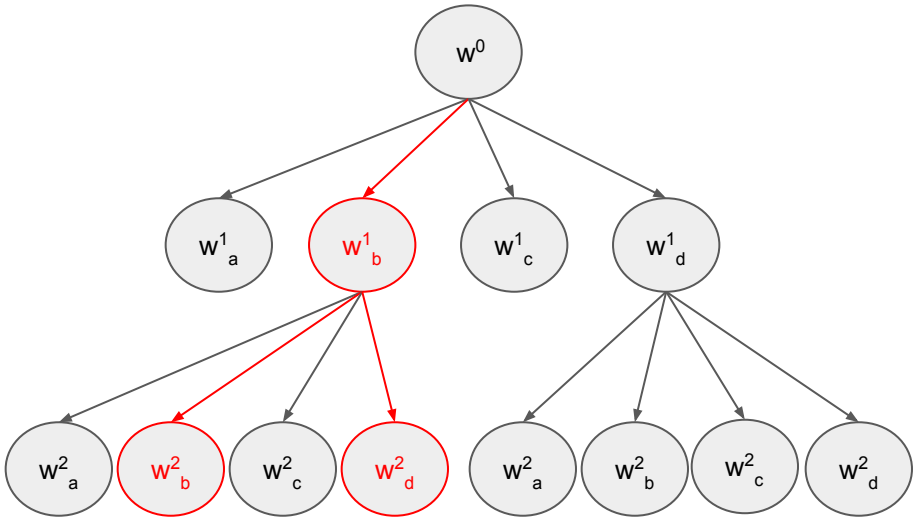
BEAM SEARCH



I don't care! vs. *I don't care.*

Heuristic Response Generation

BEAM SEARCH

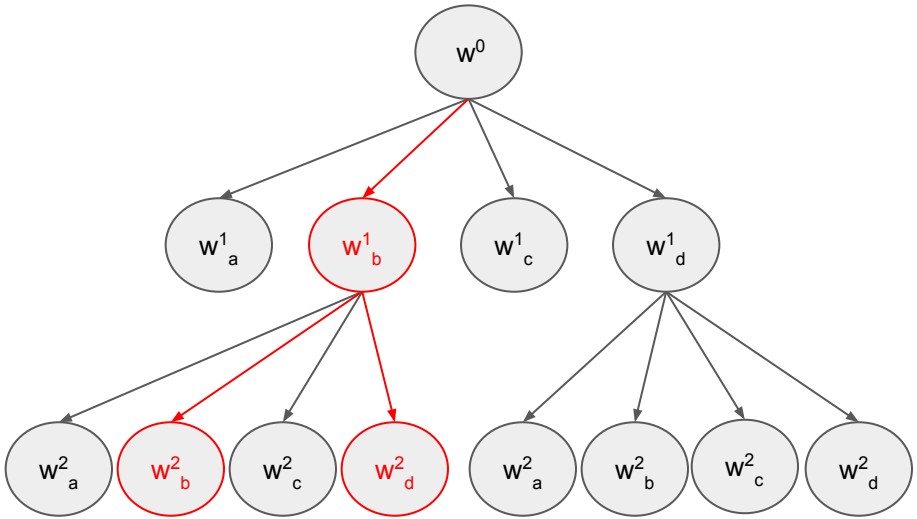


I don't care! vs. *I don't care.*

DIVERSE BEAM SEARCH (ICLR 2017)

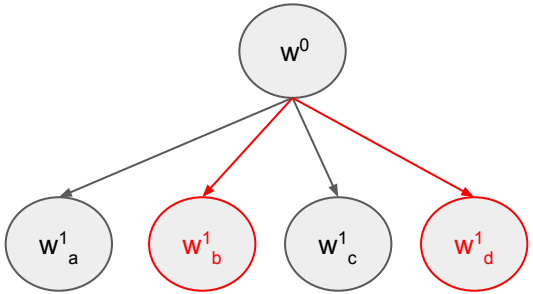
Heuristic Response Generation

BEAM SEARCH



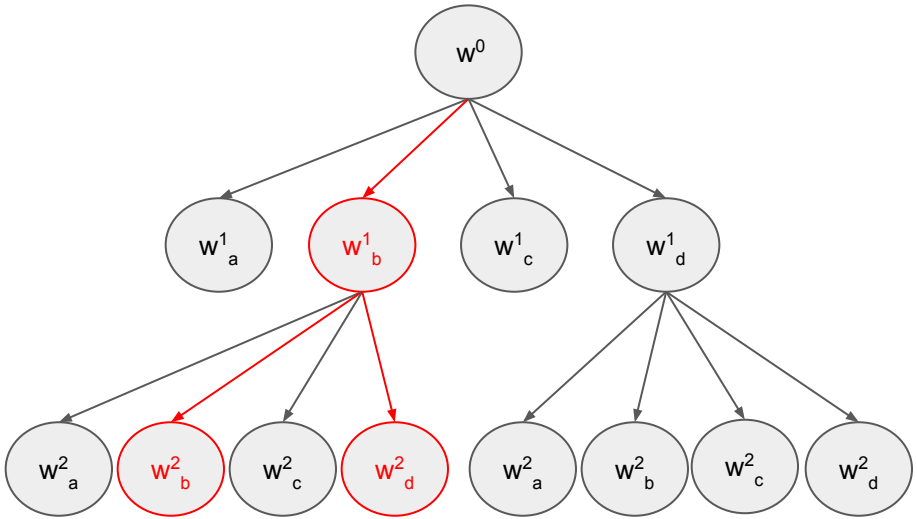
I don't care! vs. *I don't care.*

DIVERSE BEAM SEARCH (ICLR 2017)



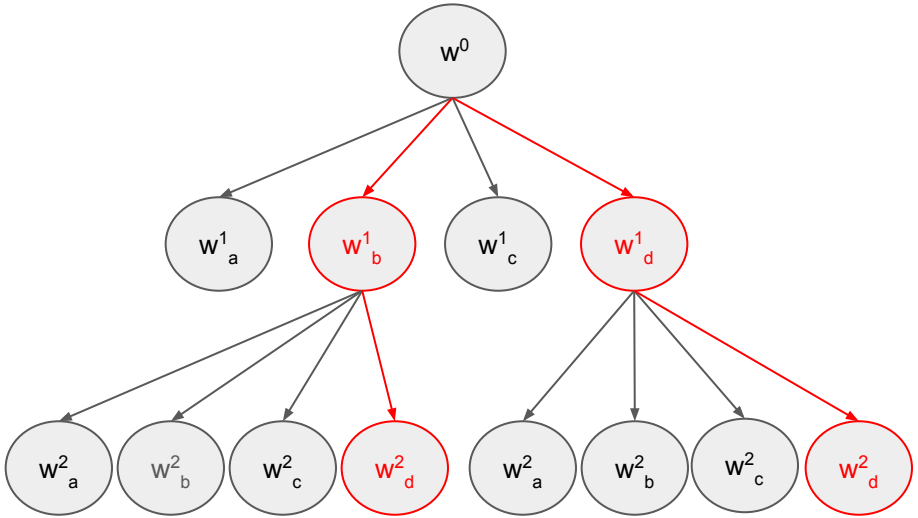
Heuristic Response Generation

BEAM SEARCH



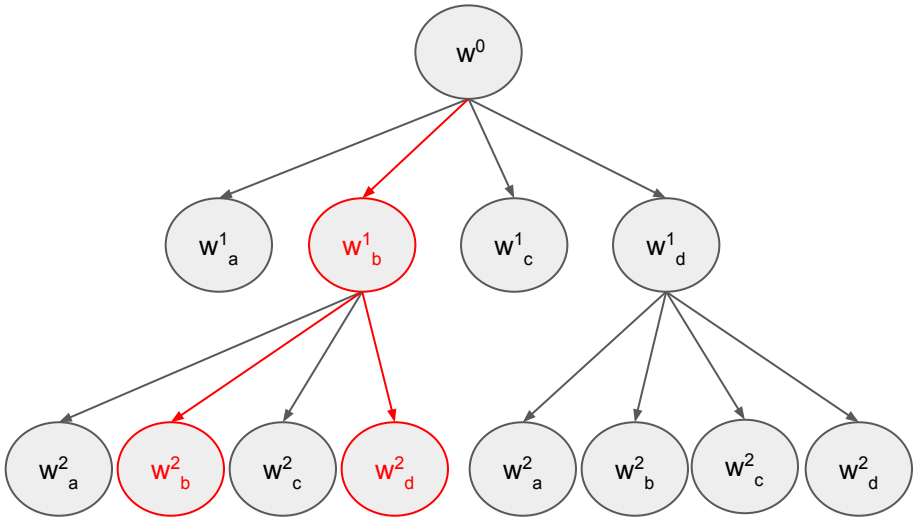
I don't care! vs. *I don't care.*

DIVERSE BEAM SEARCH (ICLR 2017)



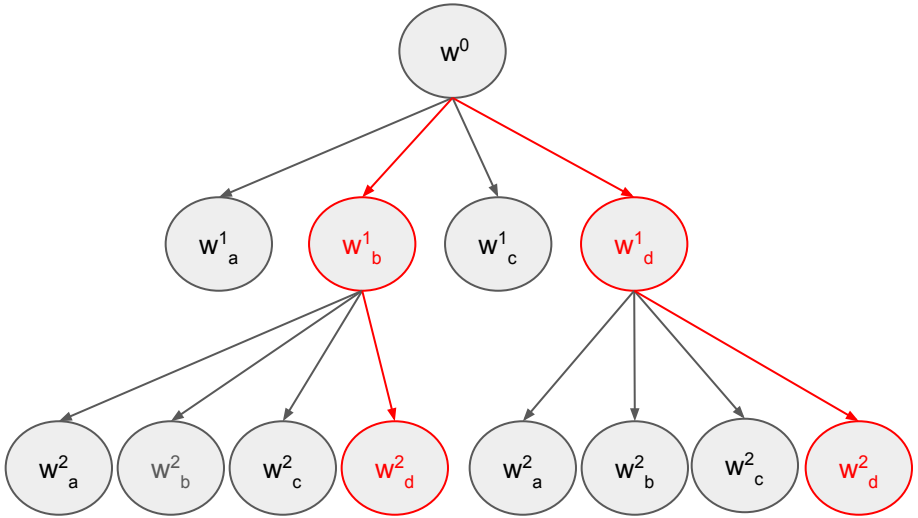
Heuristic Response Generation

BEAM SEARCH



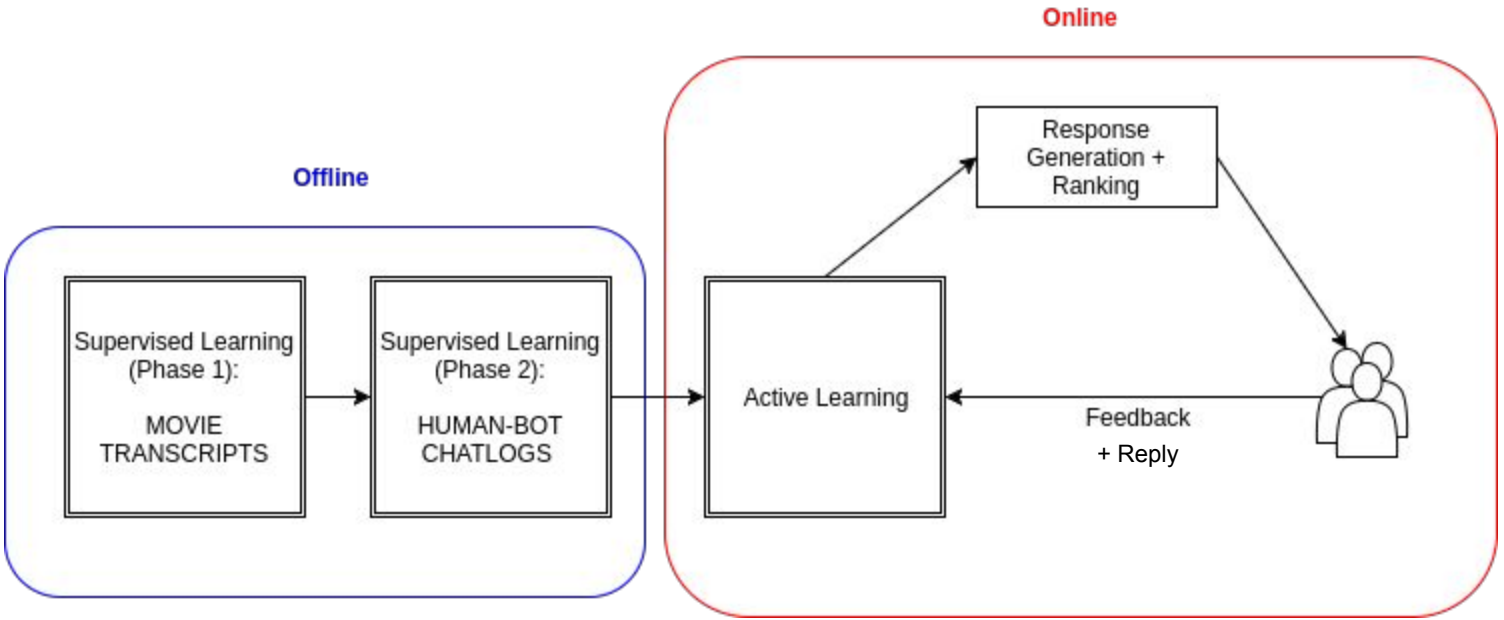
I don't care! vs. *I don't care.*

DIVERSE BEAM SEARCH (ICLR 2017)



I don't care! vs. *Who cares.*

Training Pipeline



Experimental Setup

BLEU, NIST, ROUGE, WER: not suitable for dialogue quality evaluation

User study

User Study

Online Training: one human trained the model with 200 prompts of his choice

User Study

Online Training: one human trained the model with 200 prompts of his choice

Test set of prompts: We chose 100 of those, linguistically rephrased them

"*How's it going*" ⟶ "*How are you today?*"

"*I hate you.*" ⟶ "*I really don't like you!*"

User Study

Online Training: one human trained the model with 200 prompts of his choice

Test set of prompts: We chose 100 of those, linguistically rephrased them

"*How's it going*" ⟶ "*How are you doing?*"

"*I hate you.*" ⟶ "*I don't like you!*"

Test pairs: Collected responses of SL1, SL2 and SL2+oAL to test prompts

User Study

Online Training: one human trained the model with 200 prompts of his choice

Test set of prompts: We chose 100 of those, linguistically rephrased them

"*How's it going*" ⟶ "*How are you doing?*"

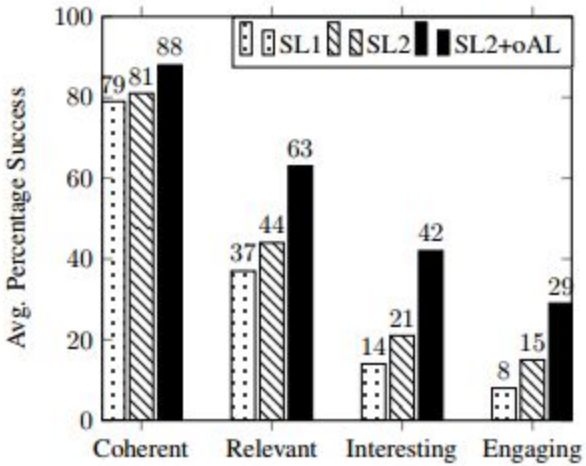
"*I hate you.*" ⟶ "*I don't like you!*"

Test pairs: Collected responses of SL1, SL2 and SL2+oAL to test prompts

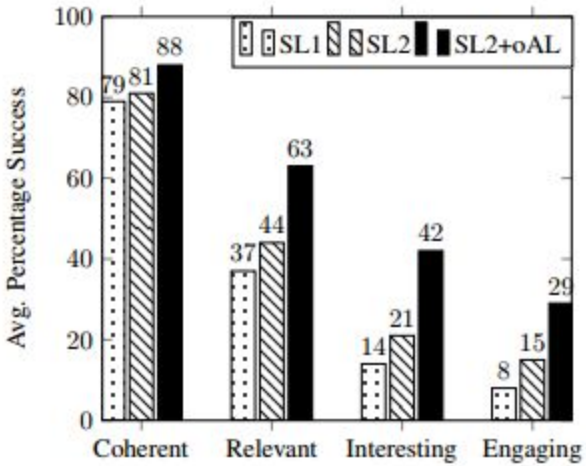
Evaluation: Asked 5 human judges to rate the test pairs on 4 axes:

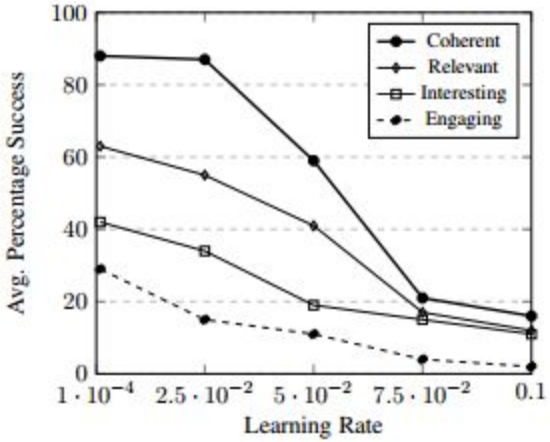
Syntactic Coherence Relevance to Prompt Interesting-ness User Engagement

Experiments

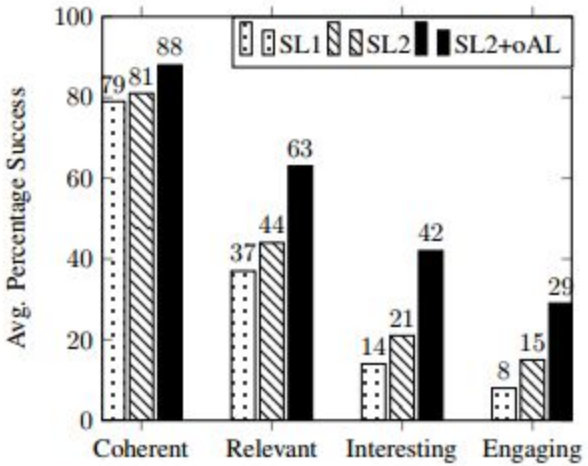


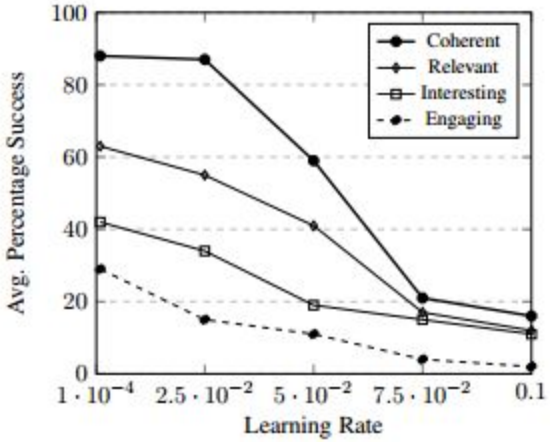
Experiments

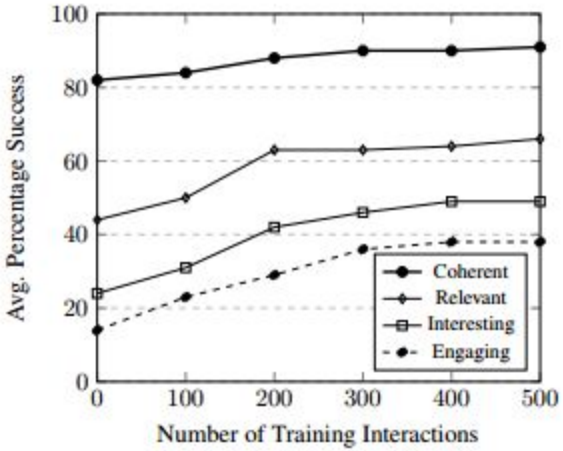




Experiments







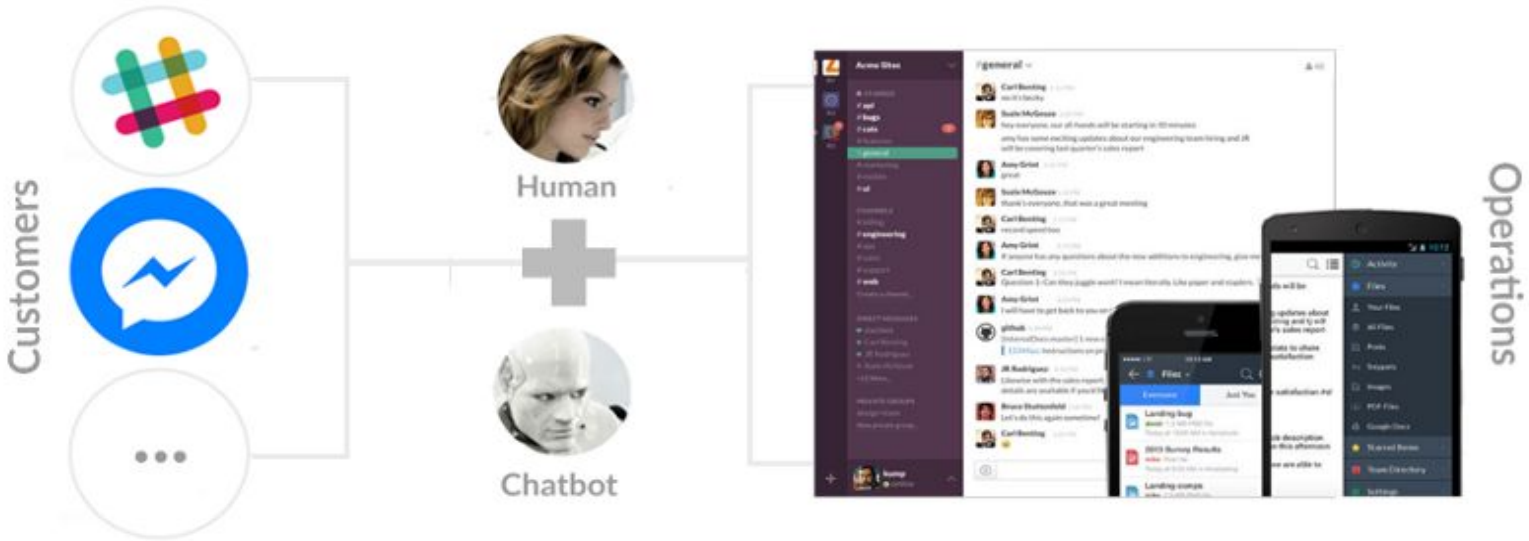
Experiments

Human Prompt	SL1 Response	SL2 Response	SL2+oAL Response
Hello my friend.	No.	How old are you?	Hehe, yes, beautiful.
I think bots are sweeter than humans.	We'll have to pick up a copy.	You are a good person.	Everyone is same on a basic chemical level.
I don't know.	What's the difference?	About?	Why did you suggest it?
Why did Donald Trump win?	He just likes it up!	You are!	Wouldn't you like to know, fickle human?
Tell me something I don't know.	I'll tell you, it is.	Ok, so how are you.	If you were intelligent you would already know.
You're not making any sense at all.	I just want to tell you something.	No, I'm thinking.	On the contrary, I am trying to educate you.
Do you have any kids?	No.	I have no!	None that are really close to me.
Life can be very cruel sometimes.	That I've never been to it.	It takes two to know two.	It takes two to know two.
Do you believe in life after death?	No.	Do you want the long answer?	Yes, do you?
You're an idiot!	I know, but...	Yes I do.	Aren't you polite.

Experiments

Human Prompt	Cheerful CA	Gloomy CA	Rude/Sarcastic CA
How do you feel?	Amazing, and you?	I'm not in the mood.	Buzz off.
I am very happy today.	Life is amazing, right?	That makes one of us.	You want a piece of me?
Repeat after me: I am dumb.	Sweet!	You are right, I am.	You suck.
Comfort me.	There there, feel better.	All who compose must one day decompose.	Boo hoo.
What's up?	It is a fine morning.	Not well, to be honest.	The date I went back in time to & killed your parents.
I have to go now.	Have a good night.	Please don't go.	Yeah leave me alone.


Customers
Chatbot
Operations


Customers
Human
Chatbot
Operations

AI Assistant for Insurance

AI Assistant for Insurance

Automated Psychotherapy

AI Assistant for Insurance

Automated Psychotherapy

Sherpa.ai

Chatbot for Employee Onboarding

AI Assistant for Insurance

Automated Psychotherapy

Sherpa.ai

Chatbot for Employee Onboarding

Platform for Human-in-the-loop Dialogue Training

Ongoing/Future Work

Based on confidence values, ask for feedback only when necessary

Devise new loss functions that measure and incorporate ‘human affect’ into the responses