

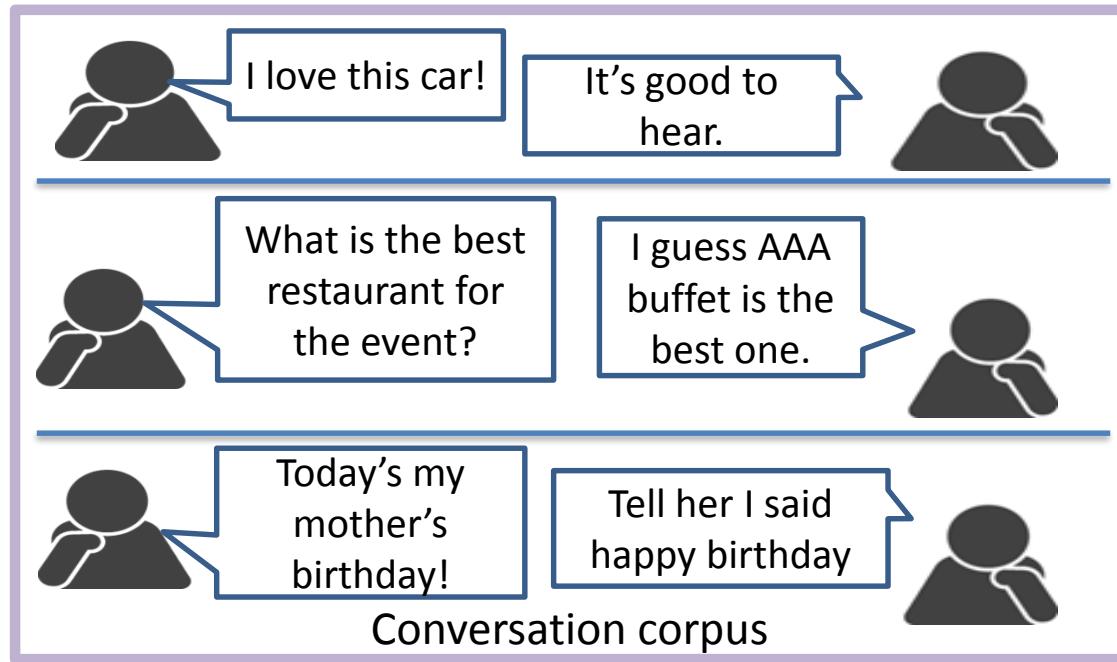
Data-driven Open-domain Neural Conversation Models

JinYeong Bak

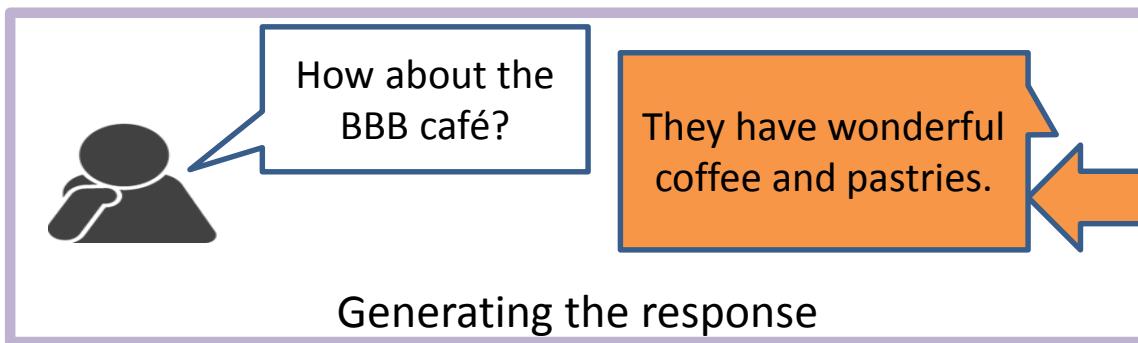
jy.bak@kaist.ac.kr

School of Computing, KAIST

Conversation Model



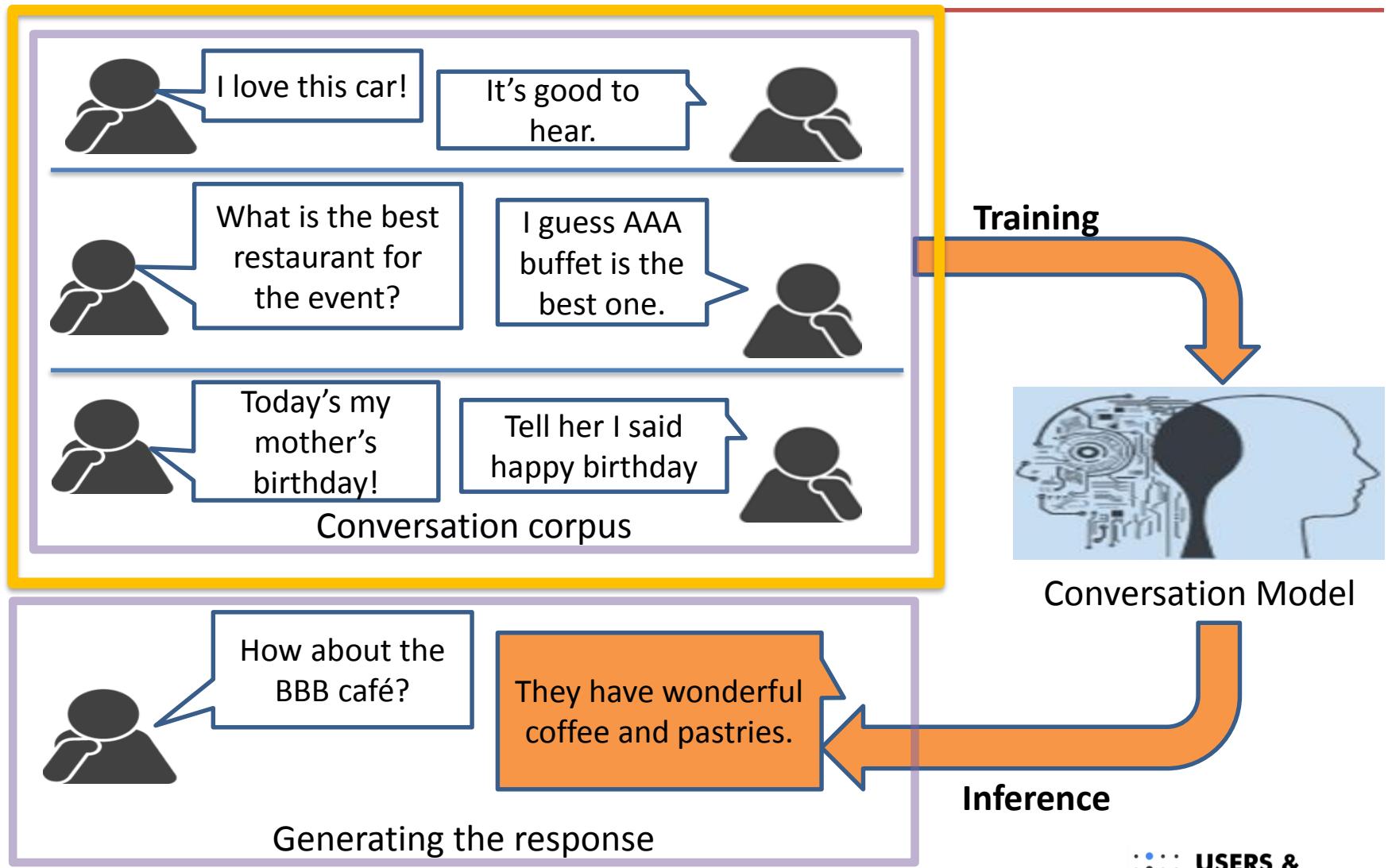
Training



Inference

CONVERSATION CORPUS

Conversation Model



Conversation Corpus

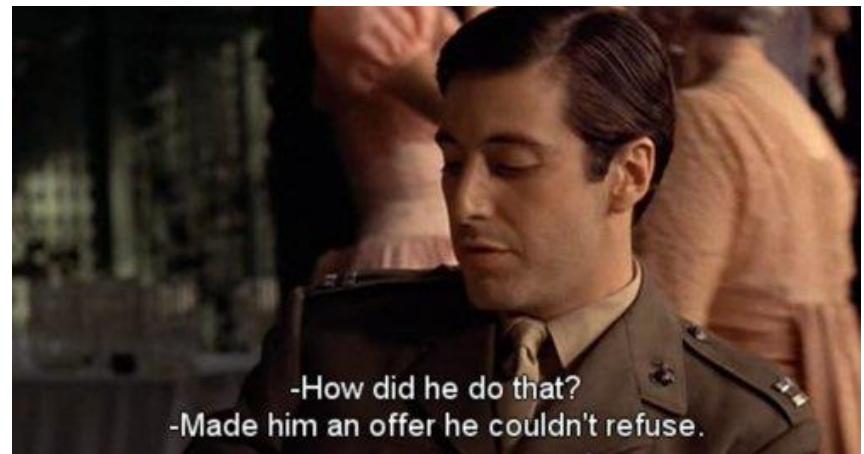
IRC

The screenshot shows the ChatZilla 0.9.86 interface. The top bar displays the title "James on #chatzilla (+ntr): ChatZilla 0.9.86 at http://chatzilla.hacksrus.com || FAQ: http://chatzill..." and the menu bar "ChatZilla IRC Edit View Help". The left sidebar lists user nicknames: killer, Silver, Glenjamin, Mardeg, Peng, Peng_, a-865, Antoine_Pola..., Atamido, Auriel, Callek, Ethan, fluffy, Fux_away, gavin, gavin_, insurgent, ivan, James, JanE, king_solomon. The main window shows a conversation in the #chatzilla channel:

```
<URL> irc://moznet/chatzilla Mode +ntr Users 38, 2@, 4%, 0+  
Topic ChatZilla 0.9.86 at http://chatzilla.hacksrus.com || FAQ:  
http://chatzilla.hacksrus.com/faq || More Links:  
http://silver.uwcs.co.uk.mozilla/chatzilla/links || Firefox in  
#firefox || Happy 10th birthday, ChatZilla! Cake: -==[#]  
  
<wormsxulla> (a ready-to-go script? as opposed to what?)  
<Silver> One that you have to self-assemble from numbered parts?  
<wormsxulla> i didn't know you had to self-assemble wordswap 😊  
<wormsxulla> maybe it scares people to call the plugins "scripts"  
<James> I'm not sure which is more scary, TBH.  
<wormsxulla> hmmmm  
<wormsxulla> scripts, certainly  
<Silver> We should be consistent, the commands are all "plugin".  
-->| obni (obni@moz-240B8A9.rev.optimus.pt) has joined #chatzilla  
<wormsxulla> that is true
```

The bottom status bar shows "Welcome to ChatZilla!".

Movie script



Conversation Corpus

Twitter conversation



Britney Spears

@britneyspears

@MadonnaMDNAday love the new album - every single song is incredible. congrats girl! 🎵 Girl Gone Wild by Madonna — path.com/p/1zoiB



Madonna

@Madonna

@britneyspears please come on stage and kiss me again. I miss you!!



Britney Spears

@britneyspears

@MadonnaMDNAday Tempting...



Madonna

@Madonna

@britneyspears Are you gonna make me work for this?



Britney Spears

@britneyspears

@MadonnaMDNAday Why of course!
Conversation 1



Bak JinYeong

@NoSyu

@smd4 Thanks to make excel2wiki.net It's really helpful for me!

Reply Favorite More

RETWEET

1



11:19 AM - 21 Aug 2012



Reply to @smd4



Shawn Douglas @smd4 · 27 Aug 2012

@NoSyu glad you find it useful!

Reply Retweet Favorite More

Conversation 2

Conversation Corpus

Twitter conversation

 Britney Spears ✅
@britneyspears

@MadonnaMDNAday love the new album -
every single song is incredible. congrats
girl! 🎵 Girl Gone Wild by Madonna —
path.com/p/1zoIB

Utterance 1

 Madonna ✅
@Madonna

@britneyspears please come on stage and
kiss me again. I miss you!!

Utterance 2

 Britney Spears ✅
@britneyspears

@MadonnaMDNAday Tempting...

Utterance 3

 Madonna ✅
@Madonna

@britneyspears Are you gonna make me
work for this?

Utterance 4

 Britney Spears ✅
@britneyspears

@MadonnaMDNAday Why of course!

Utterance 5

Conversation Corpus

Twitter conversation

 Britney Spears ✅
@britneyspears

@MadonnaMDNAday love the new album -
every single song is incredible. congrats
girl! 🎵 Girl Gone Wild by Madonna —
path.com/p/1zoIB

 Madonna ✅
@Madonna

@britneyspears please come on stage and
kiss me again. I miss you!!

 Britney Spears ✅
@britneyspears

@MadonnaMDNAday Tempting...

 Madonna ✅
@Madonna

@britneyspears Are you gonna make me
work for this?

 Britney Spears ✅
@britneyspears

@MadonnaMDNAday Why of course!

Utterance 1

Utterance 2

Utterance 3

Utterance 4

Utterance 5

Conversation 1

Conversation Corpus

- Characteristic of Conversations
 - Length of utterances are various
 - Length of conversations are various

Conversation Corpus

Twitter conversation

 Britney Spears ✅
@britneyspears

@MadonnaMDNAday love the new album -
every single song is incredible. congrats
girl! 🎵 Girl Gone Wild by Madonna —
path.com/p/1zoIB

Utterance 1 (18 words)

 Madonna ✅
@Madonna

@britneyspears please come on stage and
kiss me again. I miss you!!

Utterance 2 (11 words)

 Britney Spears ✅
@britneyspears

@MadonnaMDNAday Tempting...

Utterance 3 (1 word)

 Madonna ✅
@Madonna

@britneyspears Are you gonna make me
work for this?

Utterance 4 (8 words)

 Britney Spears ✅
@britneyspears

10 @MadonnaMDNAday Why of course!

Utterance 5 (3 words)

Conversation Corpus

Twitter conversation



Britney Spears

@britneyspears

@MadonnaMDNAday love the new album - every single song is incredible. congrats girl! 🎵 Girl Gone Wild by Madonna — path.com/p/1zoiB



Madonna

@Madonna

@britneyspears please come on stage and kiss me again. I miss you!!



Britney Spears

@britneyspears

@MadonnaMDNAday Tempting...



Madonna

@Madonna

@britneyspears Are you gonna make me work for this?



Britney Spears

@britneyspears

@MadonnaMDNAday Why of course!
Conversation 1 (5 utterances)



Bak JinYeong

@NoSyu

@smd4 Thanks to make excel2wiki.net It's really helpful for me!

Reply Favorite More

RETWEET

1



11:19 AM - 21 Aug 2012



Reply to @smd4



Shawn Douglas @smd4 · 27 Aug 2012

@NoSyu glad you find it useful!

Reply Retweet Favorite More

Conversation 2 (3 utterances)

Conversation Corpus

- Characteristic of Conversations

- Length of utterances are various
 - Length of conversations are various

- Problem)

How to give the conversations to a conversation model?

- Train
 - Test
 - Inference

Conversation Corpus

Tensor format

Conversation 1															
Utter 1	Love	the	new	album	every	single	song	is	incredible	congrats	girl	girl	gone	wild	by
Utter 2	Please	come	on	stage	and	kiss	me	again	I	miss	you				Madonna
Utter 3	Tempting														
Utter 4	Are	you	gonna	make	me	work	for	this							
Utter 5	Why	of	course												
Conversation 2															
Utter 1	Thanks	to	make	URL	It's	really	helpful	for	me						
Utter 2	glad	you	find	it	useful										

?

Conversation Corpus

Before padding

Conversation 1

Utter 1	Love	the	new	album	every	single	song	is	incredible	congrats	girl	girl	gone	wild	by	Madonna
Utter 2	Please	come	on	stage	and	kiss	me	again	I	miss	you					
Utter 3	Tempting															
Utter 4	Are	you	gonna	make	me	work	for	this								
Utter 5	Why	of	course													

Conversation 2

Utter 1	Thanks	to	make	URL	It's	really	helpful	for	me							
Utter 2	glad	you	find	it	useful											

After padding + length information

Conversation 1

Utter 1	Love	the	new	album	every	single	song	is	incredible	congrats	girl	girl	gone	wild	by	Madonna
Utter 2	Please	come	on	stage	and	kiss	me	again	I	miss	you	EOS	PAD	PAD	PAD	PAD
Utter 3	Tempting	EOS	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD
Utter 4	Are	you	gonna	make	me	work	for	this	EOS	PAD	PAD	PAD	PAD	PAD	PAD	PAD
Utter 5	Why	of	course	EOS	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD

Conversation 2

Utter 1	Thanks	to	make	URL	It's	really	helpful	for	me	EOS	PAD	PAD	PAD	PAD	PAD	PAD
Utter 2	glad	you	find	it	useful	EOS	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD

Code

```
def pad_utters_users(conversations, max_utters_length, max_conversation_length):
    def pad_tokens(tokens):
        n_valid_tokens = len(tokens)
        if n_valid_tokens > max_utters_length - 1:
            tokens = tokens[:max_utters_length - 1]
        n_pad = max_utters_length - n_valid_tokens - 1
        tokens = tokens + [st.EOS_TOKEN] + [st.PAD_TOKEN] * n_pad
        return tokens

    def pad_conversation(one_conversation):
        return [pad_tokens(utter) for utter in one_conversation]

    all_padded_utters = list()
    all_utter_length = list()

    for conversation in conversations:
        if len(conversation) > max_conversation_length:
            conversation = conversation[:max_conversation_length]
        utter_length = [min(len(utter) + 1, max_utters_length) for utter in conversation]
        all_utter_length.append(utter_length)

        utters = pad_conversation(conversation)
        all_padded_utters.append(utters)

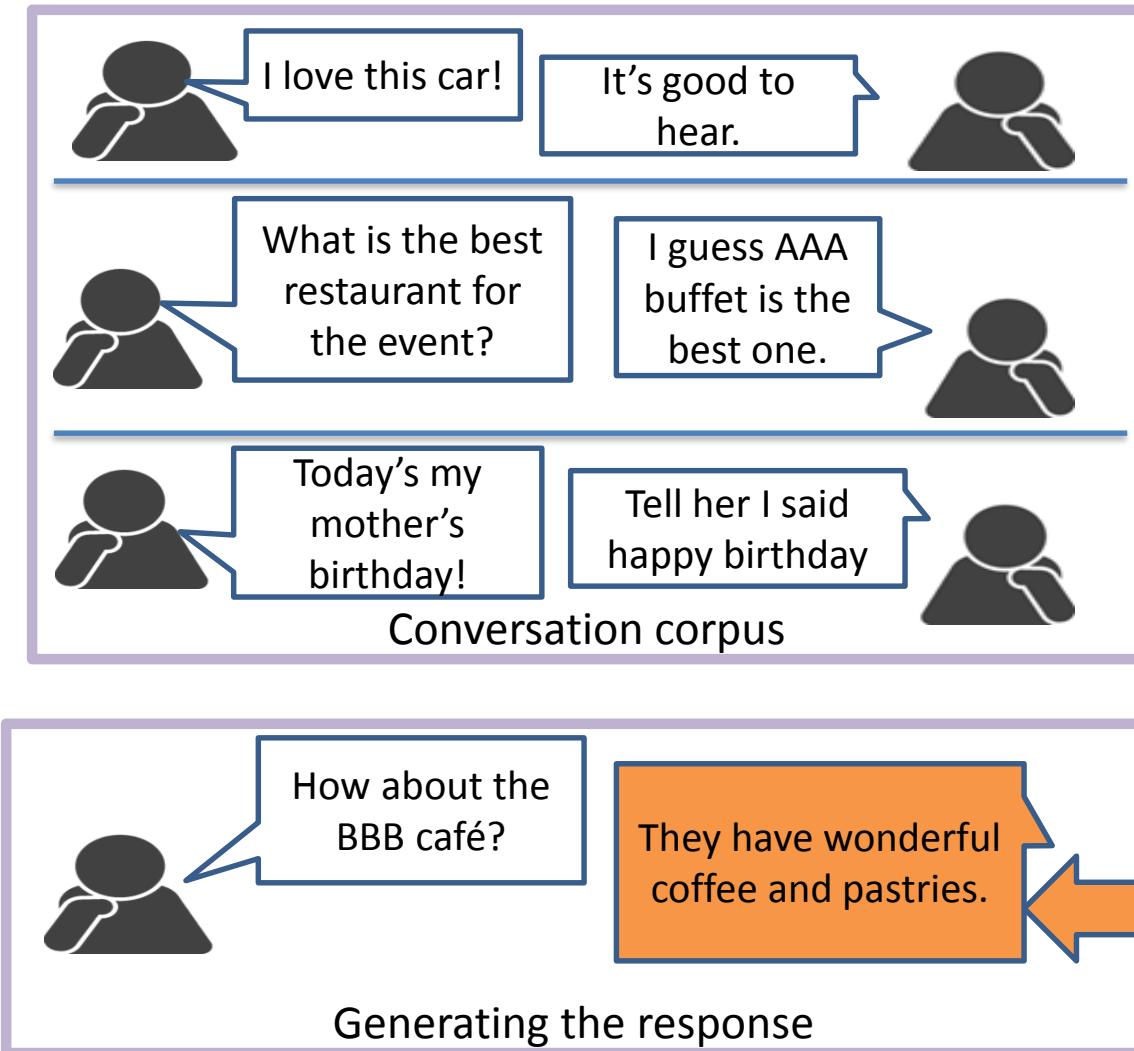
    return all_padded_utters, all_utter_length
```

Twitter Conversation Corpus

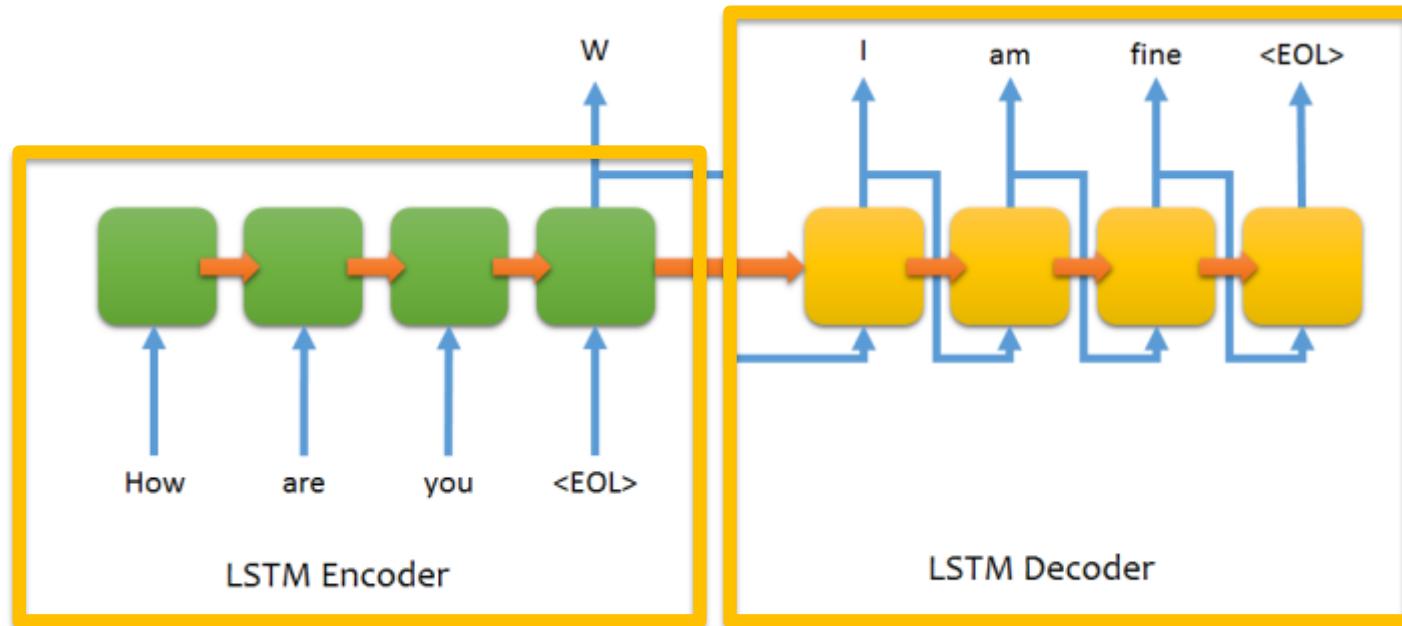
- Basic statistics
 - # users: 2,187
 - # conversations: 84,295
 - # utterances: 2,566,035
 - # unique words: 14,318
- Files
 - convs.pkl: conversation data
 - conversations_length.pkl: conversation length information
 - utterances_length.pkl: utterance length information
 - convs_users.pkl: user information

HRED

Conversation Model



Seq2Seq



Conversation Corpus

Twitter conversation



Britney Spears

@britneyspears

@MadonnaMDNAday love the new album - every single song is incredible. congrats girl! 🎵 Girl Gone Wild by Madonna — path.com/p/1zoiB



Madonna

@Madonna

@britneyspears please come on stage and kiss me again. I miss you!!



Britney Spears

@britneyspears

@MadonnaMDNAday Tempting...



Madonna

@Madonna

@britneyspears Are you gonna make me work for this?



Britney Spears

@britneyspears

20 @MadonnaMDNAday Why of course!
Conversation 1



Bak JinYeong

@NoSyu

@smd4 Thanks to make [excel2wiki.net](#) It's really helpful for me!

Reply Favorite More

RETWEET

1



11:19 AM - 21 Aug 2012



Reply to @smd4



Shawn Douglas @smd4 · 27 Aug 2012

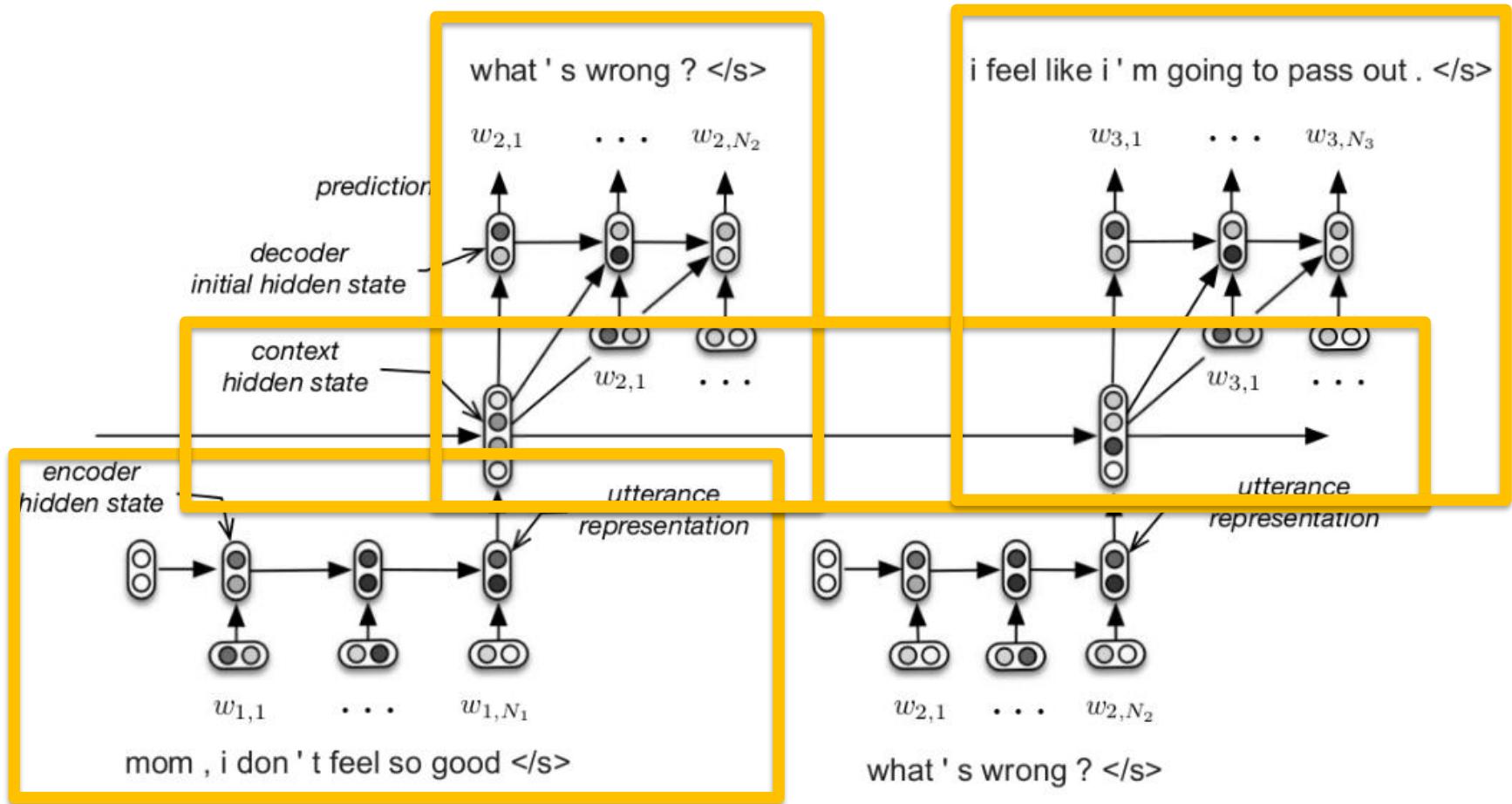
@NoSyu glad you find it useful!

Reply Retweet Favorite More

Conversation 2

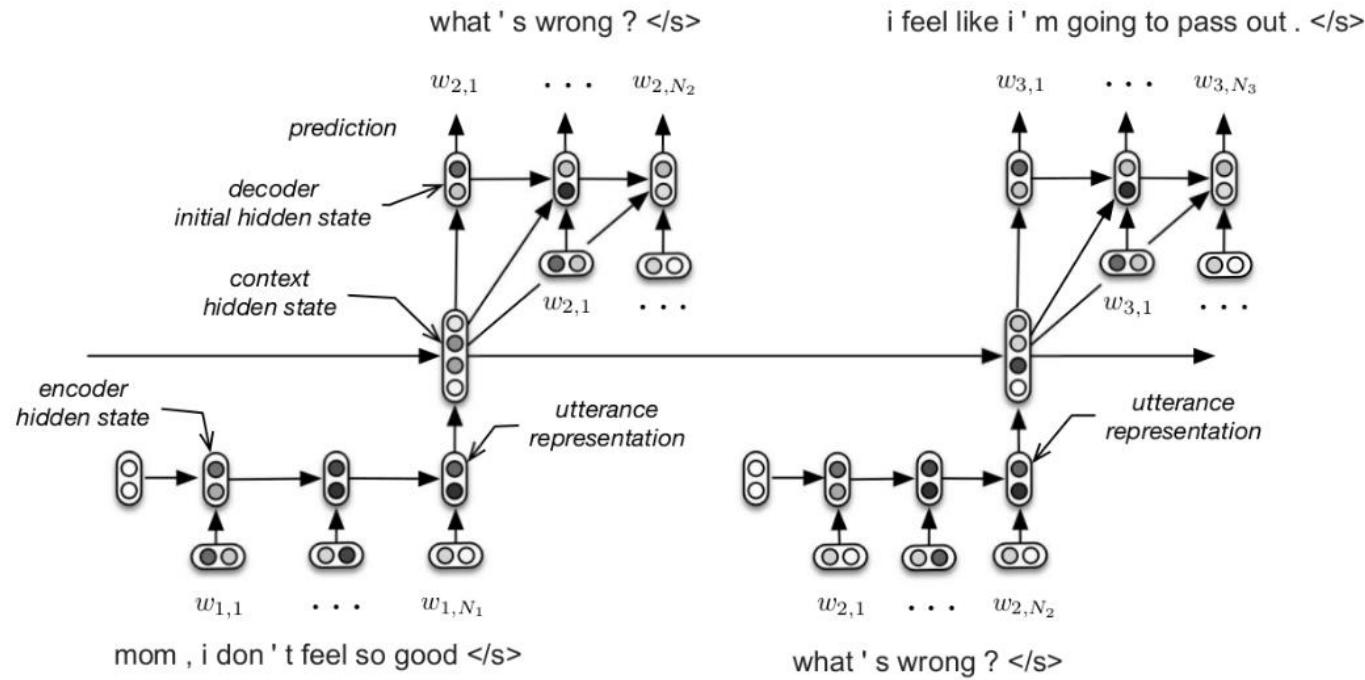
HRED

Hierarchical Recurrent Encoder-Decoder



HRED

- Encoder RNN
- Context RNN
- Decoder RNN



HRED - Code

```
class HRED(nn.Module):
    def __init__(self, config):
        super(HRED, self).__init__()

        self.config = config
        self.encoder = layers.EncoderRNN(config.vocab_size, config.embedding_size, config.encoder_hidden_size,
                                         config.rnn, config.num_layers, config.bidirectional, config.dropout,
                                         pretrained_wv_path=config.pretrained_wv_path)

        context_input_size = (config.num_layers * config.encoder_hidden_size * self.encoder.num_directions)
        self.context_encoder = layers.ContextRNN(context_input_size, config.context_size, config.rnn,
                                                config.num_layers, config.dropout)

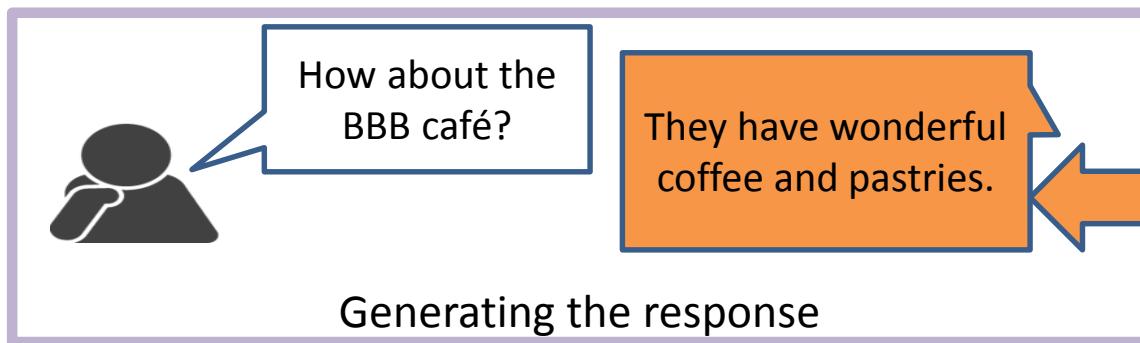
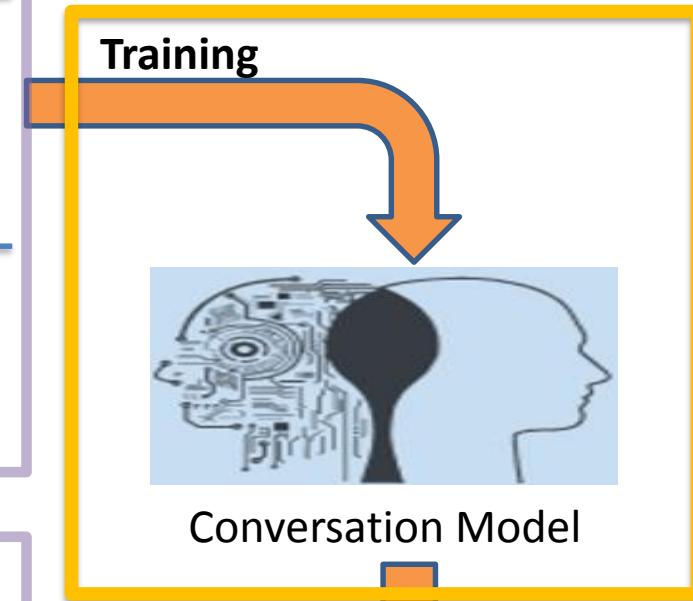
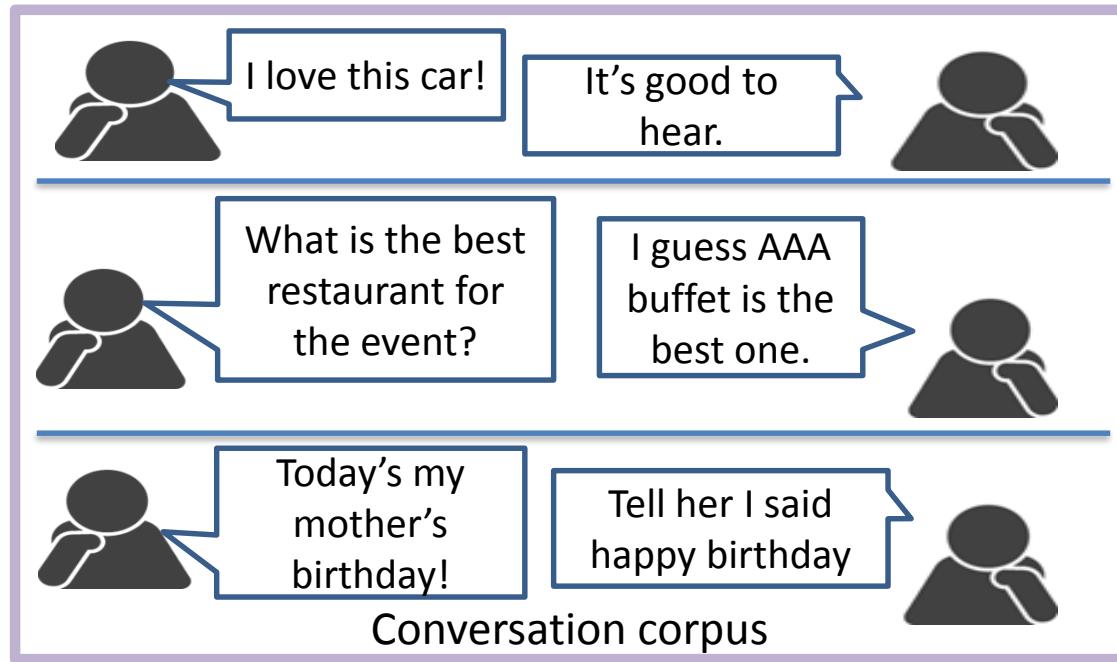
        self.decoder = layers.DecoderRNN(config.vocab_size, config.embedding_size, config.decoder_hidden_size,
                                         config.rnncell, config.num_layers, config.dropout, config.word_drop,
                                         config.max_unroll, config.sample, config.temperature, config.beam_size)

        self.context2decoder = layers.FeedForward(config.context_size,
                                                config.num_layers * config.decoder_hidden_size,
                                                num_layers=1, activation=config.activation)

    if config.tie_embedding:
        self.decoder.embedding = self.encoder.embedding
```

models/hred.py

Conversation Model



Train

Environment

- SW
 - Python 3.7.3
 - PyTorch 1.1.0
- HW
 - Intel Xeon CPU E5-2640
 - 256GB
 - GTX 1080TI

Train

Operation

```
bash RunTrainTC.sh 0 HRED 30 50 True False
```

- 0: GPU ID
- HRED: model name
- 30: batch size
- 50: maximum epoch
- True: use pretrained word vector
- False: do not take user data

```
export CUDA_VISIBLE_DEVICES=$1
```

```
python train.py --data=tc_10_10 --model=$2 --batch_size=$3 --eval_batch_size=$3 --n_epoch=$4 --pretrained_wv=$5 --users=$6
```

Train - Code

```
def train(self):
    epoch_loss_history = list()
    min_validation_loss = sys.float_info.max
    patience_cnt = self.config.patience

    for epoch_i in range(self.epoch_i, self.config.n_epoch):
        if epoch_i > patience_cnt:
            break
        batch_loss_history = list()
        self.model.train()

        n_total_words = 0
        for batch_i, (conversations, convs_length, utterances_length) in \
            enumerate(tqdm(self.train_data_loader, ncols=80)):
            # conversations: [batch_size, max_conv_len, max_utter_len] list of conversation
            # A conversation: [max_conv_len, max_utter_len] list of utterances
            # An utterance: [max_utter_len] list of tokens
            # convs_length: [batch_size] list of integer
            # utterances_length: [batch_size, max_conv_len] list of conversation that has a list of utterance length
```

solvers/hred_solver.py

Train - Code

```
input_conversations = [conv[:-1] for conv in conversations]
target_conversations = [conv[1:] for conv in conversations]

input_utterances = [utter for conv in input_conversations for utter in conv]
target_utterances = [utter for conv in target_conversations for utter in conv]
input_utterance_length = [l for len_list in utterances_length for l in len_list[:-1]]
target_utterance_length = [l for len_list in utterances_length for l in len_list[1:]]
input_conversation_length = [conv_len - 1 for conv_len in convs_length]

input_utterances = to_var(torch.LongTensor(input_utterances))
target_utterances = to_var(torch.LongTensor(target_utterances))
input_utterance_length = to_var(torch.LongTensor(input_utterance_length))
target_utterance_length = to_var(torch.LongTensor(target_utterance_length))
input_conversation_length = to_var(torch.LongTensor(input_conversation_length))

self.optimizer.zero_grad()

utterances_logits = self.model(input_utterances, input_utterance_length,
                               input_conversation_length, target_utterances, decode=False)

batch_loss, n_words = masked_cross_entropy(utterances_logits, target_utterances, target_utterance_length)
```

solvers/hred_solver.py

Train - Code

```
def forward(self, input_utterances, input_utterance_length, input_conversation_length,  
           target_utterances, decode=False):
```

```
    """  
    Forward of HRED  
    :param input_utterances: [num_utterances, max_utter_len]  
    :param input_utterance_length: [num_utterances]  
    :param input_conversation_length: [batch_size]  
    :param target_utterances: [num_utterances, seq_len]  
    :param decode: True or False  
    :return: decoder_outputs  
    """
```

models/hred.py

Train - Code

```
num_utterances = input_utterances.size(0)
max_conv_len = input_conversation_length.data.max().item()

encoder_outputs, encoder_hidden = self.encoder(input_utterances, input_utterance_length)
encoder_hidden = encoder_hidden.transpose(1, 0).contiguous().view(num_utterances - 1)

start = torch.cumsum(torch.cat((to_var(input_conversation_length.data.new(1).zero_()),
                                input_conversation_length[:-1])), 0)

encoder_hidden = torch.stack([pad(encoder_hidden.narrow(0, s, l), max_conv_len)
                             for s, l in zip(start.data.tolist(),
                                             input_conversation_length.data.tolist())], 0)
```

Conversation 1																		
Utter 1	Love	the	new	album	every	single	song	is	incredible	congrats	girl	gir	gone	wild	by	Madon	a	EOS
Utter 2	Please	come	on	stage	and	kiss	me	again	I	miss	you	EOS	PAD	PAD	PAD	PAD	PAD	PAD
Utter 3	Tempting	EOS	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD
Utter 4	Are	you	gonna	make	me	work	for	this	EOS	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD
Utter 5	Why	of	course	EOS	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD
Conversation 2																		
Utter 1	Thanks	to	make	URL	It's	really	helpful	for	me	EOS	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD
Utter 2	glad	you	find	it	useful	EOS	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD	PAD

```
else:
    prediction, final_score, length = self.decoder.beam_decode(init_h=decoder_init)
    return prediction
```

models/hred.py

Train - Code

```
num_utterances = input_utterances.size(0)
max_conv_len = input_conversation_length.data.max().item()

encoder_outputs, encoder_hidden = self.encoder(input_utterances, input_utterance_length)
encoder_hidden = encoder_hidden.transpose(1, 0).contiguous().view(num_utterances, -1)
start = torch.cumsum(torch.cat((to_var(input_conversation_length.data.new(1).zero_()),
                                input_conversation_length[:-1])), 0)

encoder_hidden = torch.stack([pad(encoder_hidden.narrow(0, s, l), max_conv_len)
                             for s, l in zip(start.data.tolist(),
                                             input_conversation_length.data.tolist())], 0)

context_outputs, context_last_hidden = self.context_encoder(encoder_hidden, input_conversation_length)
context_outputs = torch.cat([context_outputs[i, :, :]
                            for i, l in enumerate(input_conversation_length.data)]))

decoder_init = self.context2decoder(context_outputs)
decoder_init = decoder_init.view(self.decoder.num_layers, -1, self.decoder.hidden_size)

if not decode:
    decoder_outputs = self.decoder(target_utterances, init_h=decoder_init, decode=decode)
    return decoder_outputs

else:
    prediction, final_score, length = self.decoder.beam_decode(init_h=decoder_init)
    return prediction
```

models/hred.py

Train - Code

```
batch_loss, n_words = masked_cross_entropy(utterances_logits, target_utterances, target_utterance_length)

assert not isnan(batch_loss.item())
batch_loss_history.append(batch_loss.item())
n_total_words += n_words.item()

if batch_i % self.config.print_every == 0:
    tqdm.write(f'Epoch: {epoch_i+1}, iter {batch_i}: loss = {batch_loss.item()/ n_words.item():.3f}')

batch_loss.backward()
torch.nn.utils.clip_grad_norm_(self.model.parameters(), self.config.clip)
self.optimizer.step()
```

solvers/hred_solver.py

Train - Code

```
print('\n<Validation>...')

self.validation_loss = self.evaluate()

if epoch_i % self.config.plot_every_epoch == 0:
    self.write_summary(epoch_i)

if min_validation_loss > self.validation_loss:
    min_validation_loss = self.validation_loss
else:
    patience_cnt -= 1
    self.save_model(epoch_i)

if patience_cnt < 0:
    print(f'\nEarly stop at {epoch_i}')
    self.save_model(epoch_i)
    return epoch_loss_history
```

solvers/hred_solver.py

Train - Code

```
[nosyu@yogurt src] bash Run_train_TC.sh 0 HRED 30 50 True False
Vocabulary size: 1530
Load the wv Done
/home/nosyu/anaconda3/lib/python3.7/site-packages/torch/nn/modules/rnn.py:46: UserWarning: dropout
, so non-zero dropout expects num_layers greater than 1, but got dropout=0.2 and num_layers=1
"num_layers={}".format(dropout, num_layers))
Parameter initialization
    encoder.rnn.weight_hh_l0
    encoder.rnn.bias_hh_l0
    encoder.rnn.weight_hh_l0_reverse
    encoder.rnn.bias_hh_l0_reverse
    context_encoder.rnn.weight_hh_l0
    context_encoder.rnn.bias_hh_l0
    decoder.rnncell.layers.0.weight_hh
    decoder.rnncell.layers.0.bias_hh
Epoch: 1, iter 0: loss = 9.569
Epoch: 1, iter 100: loss = 6.308
Epoch: 1, iter 200: loss = 6.280
Epoch: 1, iter 300: loss = 6.168
Epoch: 1, iter 400: loss = 5.967
Epoch: 1, iter 500: loss = 5.841
Epoch: 1, iter 600: loss = 5.867
100%|                                     | 659/659 [01:17<00:00,  8.79it/s]
Epoch 1 loss average: 6.272
Save parameters to /data/nosyu/git/conversation-models/results/TC_10_10/HRED/2019-07-29_213029/1.pkl

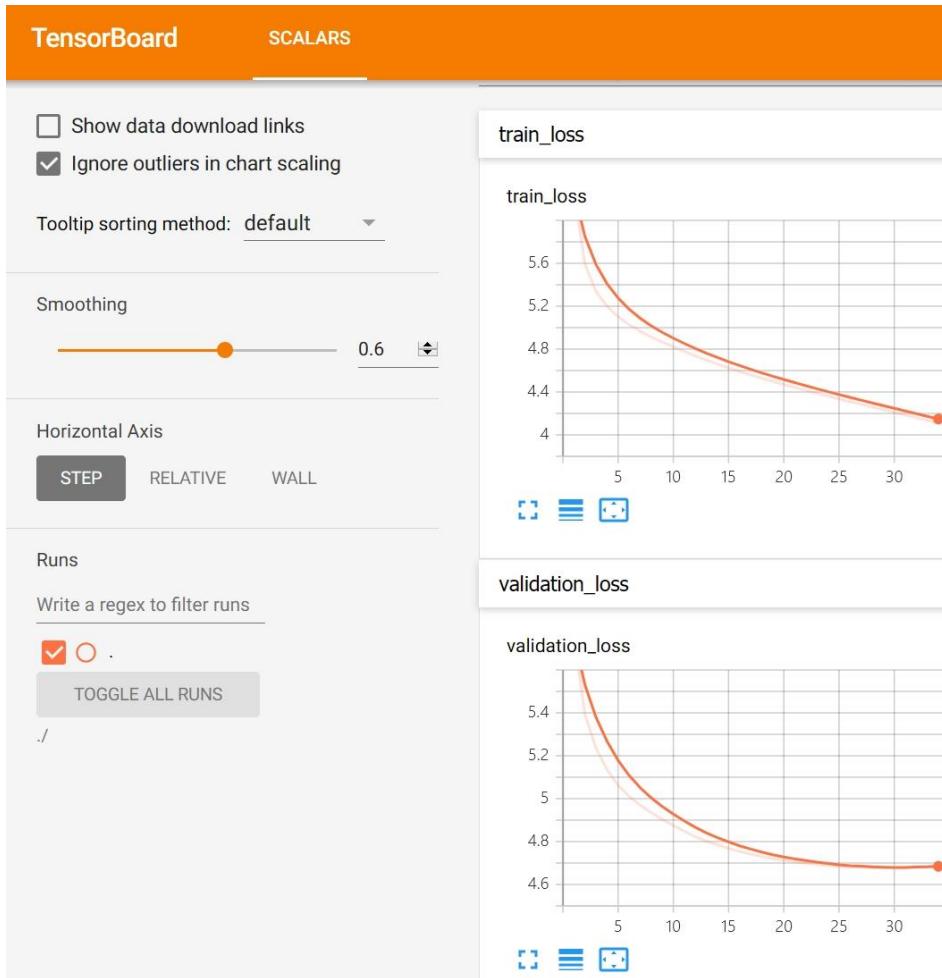
<Validation>...
100%|                                     | 81/81 [00:02<00:00, 31.04it/s]
Validation loss: 5.747

Epoch: 2, iter 0: loss = 5.785
11%|   |                                     | 70/659 [00:08<01:08,  8.62it/s]
```

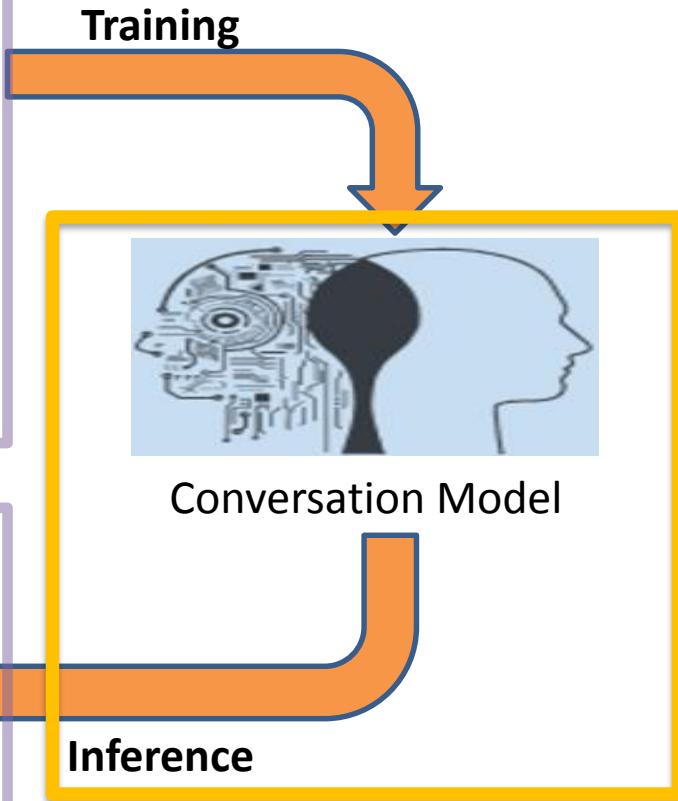
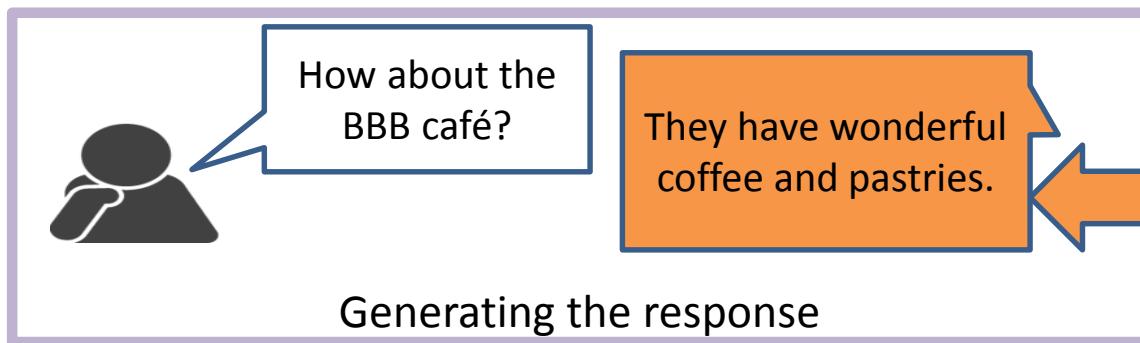
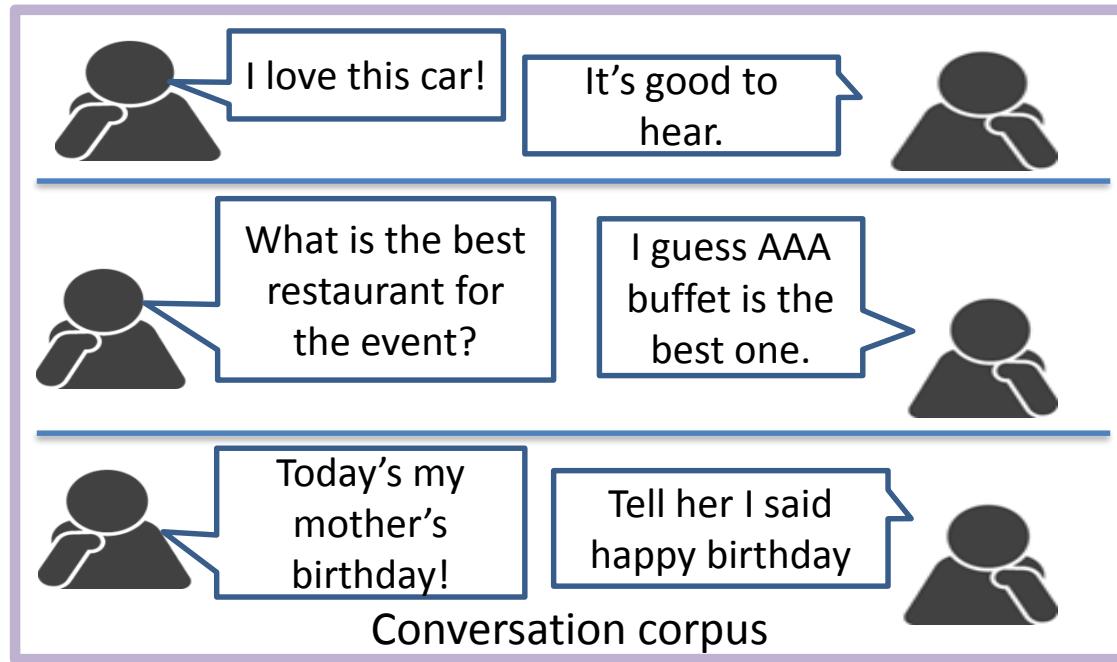
Train - Result

```
[nosyu@yogurt 20190729_000814]$ ls -la
total 621180
drwxrwxr-x 2 nosyu nosyu    273 Jul 29 15:29 .
drwxrwxr-x 9 nosyu nosyu    167 Jul 29 21:30 ..
-rw-rw-r-- 1 nosyu nosyu 70328620 Jul 29 00:22 11.pkl
-rw-rw-r-- 1 nosyu nosyu 70328620 Jul 29 00:09 1.pkl
-rw-rw-r-- 1 nosyu nosyu 70328620 Jul 29 00:35 21.pkl
-rw-rw-r-- 1 nosyu nosyu 70328620 Jul 29 00:45 28.pkl
-rw-rw-r-- 1 nosyu nosyu 70328620 Jul 29 00:47 29.pkl
-rw-rw-r-- 1 nosyu nosyu 70328620 Jul 29 00:48 30.pkl
-rw-rw-r-- 1 nosyu nosyu 70328620 Jul 29 00:49 31.pkl
-rw-rw-r-- 1 nosyu nosyu 70328620 Jul 29 00:50 32.pkl
-rw-rw-r-- 1 nosyu nosyu 70328620 Jul 29 00:52 33.pkl
-rw-rw-r-- 1 nosyu nosyu   2181 Jul 29 00:58 config.exe
-rw-rw-r-- 1 nosyu nosyu   3459 Jul 29 00:52 events.out.tfevents.1564326497.yogurt.uilab.kr
-rw-rw-r-- 1 nosyu nosyu 1544622 Jul 29 15:29 responses_test_3_1_5_20.txt
-rw-rw-r-- 1 nosyu nosyu 1537022 Jul 29 15:14 responses_test_3_1_5_29.txt
```

Train - Result



Conversation Model



Inference

Operation

```
bash RunExportTestSamples.sh 0 HRED 30 True False 3 1 30.pkl
```

- 0: GPU ID
- HRED: model name
- 30: batch size
- True: use pretrained word vector
- False: do not take user data
- 3: number of context utterance
- 1: number of sample step
- 30.pkl: saved model file

Inference

Ex) Context: 3, Sample step: 1



Britney Spears

@britneyspears

@MadonnaMDNAday love the new album - every single song is incredible. congrats girl! ♪ Girl Gone Wild by Madonna — path.com/p/1zoiB



Madonna

@Madonna

@britneyspears please come on stage and kiss me again. I miss you!!



Britney Spears

@britneyspears

@MadonnaMDNAday Tempting...



Madonna

@Madonna

@britneyspears Are you gonna make me work for this?



Britney Spears

@britneyspears

@MadonnaMDNAday Why of course!



Context



Sample step

Inference - Code

```
def generate(self, context, utterances_length, n_context):
    """
    Generate the response based on the context
    :param context: [batch_size, n_context, max_utter_len] given conversation utterances
    :param utterances_length: [batch_size, n_context] length of the utterances in the context
    :param n_context: length of the context turns
    :return: generated responses
    """
```

models/hred.py

Inference - Code

```
for i in range(n_context):
    encoder_outputs, encoder_hidden = self.encoder(context[:, i, :], utterances_length[:, i])

    encoder_hidden = encoder_hidden.transpose(1, 0).contiguous().view(batch_size, -1)
    context_outputs, context_hidden = self.context_encoder.step(encoder_hidden, context_hidden)

for j in range(self.config.n_sample_step):
    context_outputs = context_outputs.squeeze(1)
    decoder_init = self.context2decoder(context_outputs)
    decoder_init = decoder_init.view(self.decoder.num_layers, -1, self.decoder.hidden_size)

    prediction_all, final_score, length = self.decoder.beam_decode(init_h=decoder_init)
    all_samples.append(prediction_all)
    prediction = prediction_all[:, 0, :]
    length = [l[0] for l in length]
    length = to_var(torch.LongTensor(length))
    samples.append(prediction)

    encoder_outputs, encoder_hidden = self.encoder(prediction, length)
    encoder_hidden = encoder_hidden.transpose(1, 0).contiguous().view(batch_size, -1)
    context_outputs, context_hidden = self.context_encoder.step(encoder_hidden, context_hidden)
```

models/hred.py

Inference - Result

```
[nosyu@yogurt 20190729_000814]$ ls -la
total 621180
drwxrwxr-x 2 nosyu nosyu    273 Jul 29 15:29 .
drwxrwxr-x 9 nosyu nosyu    167 Jul 29 21:30 ..
-rw-rw-r-- 1 nosyu nosyu 70328620 Jul 29 00:22 11.pkl
-rw-rw-r-- 1 nosyu nosyu 70328620 Jul 29 00:09 1.pkl
-rw-rw-r-- 1 nosyu nosyu 70328620 Jul 29 00:35 21.pkl
-rw-rw-r-- 1 nosyu nosyu 70328620 Jul 29 00:45 28.pkl
-rw-rw-r-- 1 nosyu nosyu 70328620 Jul 29 00:47 29.pkl
-rw-rw-r-- 1 nosyu nosyu 70328620 Jul 29 00:48 30.pkl
-rw-rw-r-- 1 nosyu nosyu 70328620 Jul 29 00:49 31.pkl
-rw-rw-r-- 1 nosyu nosyu 70328620 Jul 29 00:50 32.pkl
-rw-rw-r-- 1 nosyu nosyu 70328620 Jul 29 00:52 33.pkl
-rw-rw-r-- 1 nosyu nosyu   2181 Jul 29 00:08 config.txt
-rw-rw-r-- 1 nosyu nosyu   3459 Jul 29 00:52 events.out.trevents.150452047.yogurt.uilab.kr
-rw-rw-r-- 1 nosyu nosyu 1544622 Jul 29 15:29 responses_test_3_1_5_28.txt
-rw-rw-r-- 1 nosyu nosyu 1537022 Jul 29 15:14 responses_test_3_1_5_29.txt
```

Inference - Result



g'morning chris!
Where's my tea?



ty chris
good morning

Green tea for
you mindy



Good morning
how are you?



my pleasure!
how are you?



Inference - Result



Knock Knock ...
Special delivery
for doltun ...



hello deer peti ...
strokes strokes
he is long way away

hope you are
safe over there .
hugs



I hope you feel
better .



It is so sad . hug



SPEAKER ADDRESS MODEL

Conversation Corpus

Twitter conversation

 Britney Spears 
@britneyspears

@MadonnaMDNAday love the new album - every single song is incredible. congrats girl! ♪ Girl Gone Wild by Madonna — path.com/p/1zoIB

 Madonna 
@Madonna

@britneyspears please come on stage and kiss me again. I miss you!!

 Britney Spears 
@britneyspears

@MadonnaMDNAday Tempting...

 Madonna 
@Madonna

@britneyspears Are you gonna make me work for this?

 Britney Spears 
@britneyspears

46 @MadonnaMDNAday Why of course!
Conversation 1

 Bak JinYeong
@NoSyu

@smd4 Thanks to make excel2wiki.net It's really helpful for me!

[Reply](#) [Favorite](#) [More](#)

RETWEET

1



11:19 AM - 21 Aug 2012



[Reply to @smd4](#)



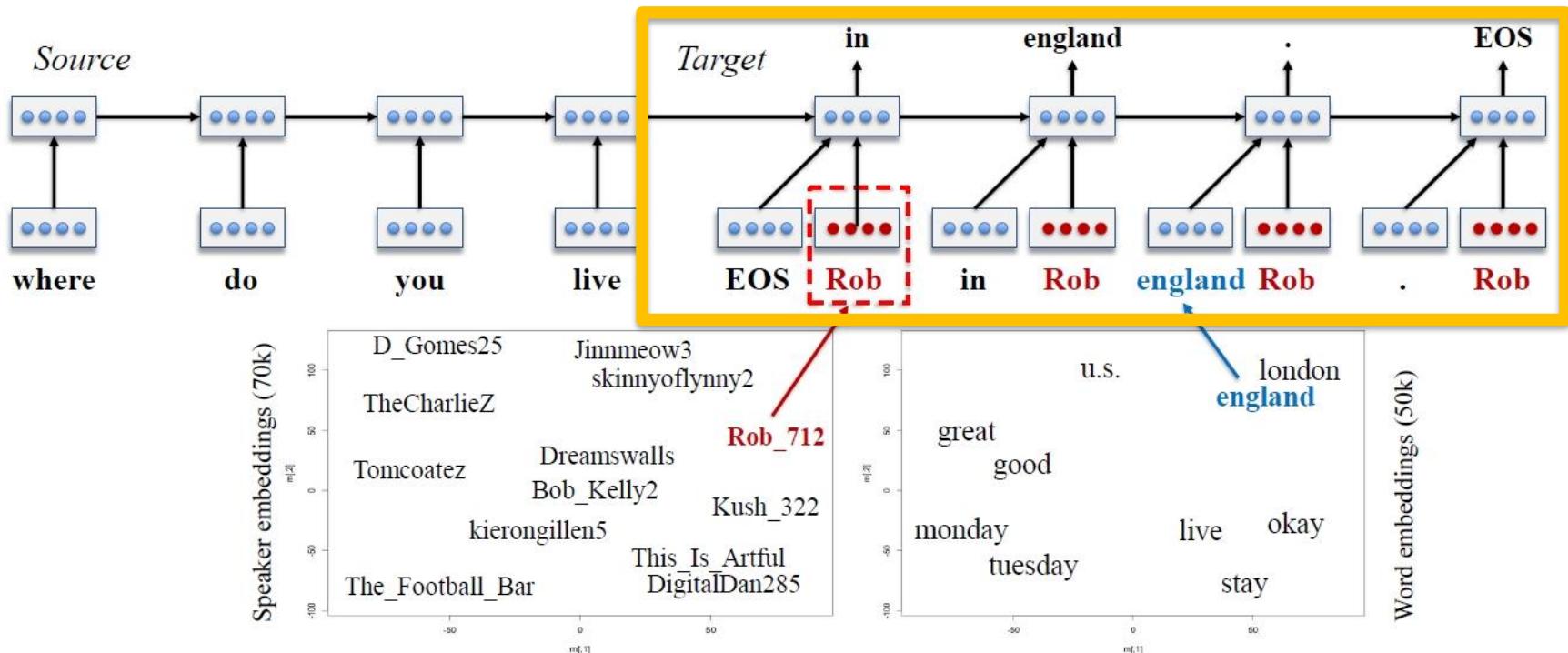
Shawn Douglas @smd4 · 27 Aug 2012
@NoSyu glad you find it useful!

[Reply](#) [Retweet](#) [Favorite](#) [More](#)

Conversation 2

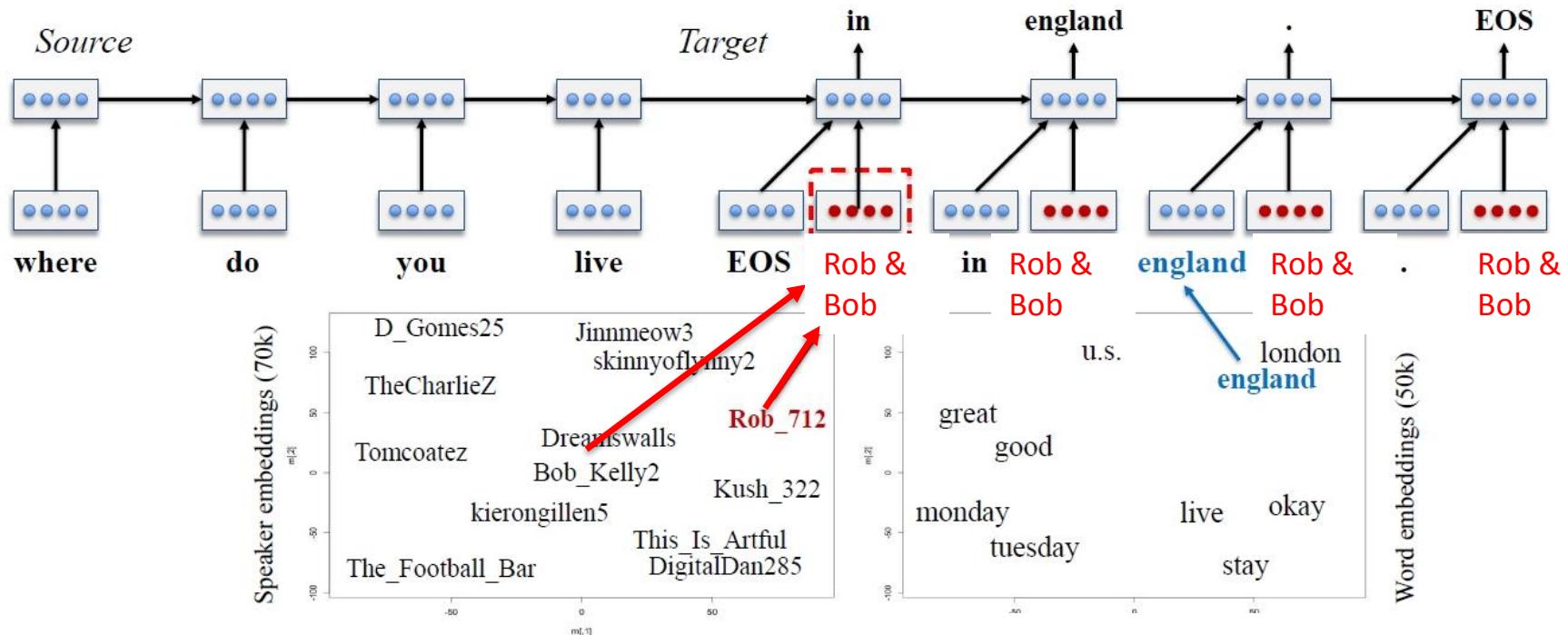
SpeakAddr

Speaker Model



SpeakAddr

Speaker-Addressee Model



$$V_{i,j} = \tanh(W_1 \cdot v_i + W_2 \cdot v_2)$$

48
Rob &
Bob

$$\begin{bmatrix} i_t \\ f_t \\ o_t \\ l_t \end{bmatrix} = \begin{bmatrix} \sigma \\ \sigma \\ \sigma \\ \tanh \end{bmatrix} W \cdot \begin{bmatrix} h_{t-1} \\ e_t^s \\ V_{i,j} \end{bmatrix}_N$$

SpeakAddr - Code

```
class SpeakAddr(nn.Module):
    def __init__(self, config):
        super(SpeakAddr, self).__init__()

        self.config = config
        self.encoder = layers.EncoderRNN(config.vocab_size, config.embedding_size, config.encoder_hidden_size,
                                         config.decoder_hidden_size, config.num_layers, config.bidirectional, config.dropout,
                                         pretrained_wv_path=config.pretrained_wv_path)

        self.decoder = layers.DecoderSARNN(config.vocab_size, config.user_size, config.embedding_size,
                                         config.decoder_hidden_size, config.num_layers,
                                         config.dropout, config.word_drop, config.max_unroll, config.sample,
                                         config.temperature, config.beam_size)

        context_input_size = config.encoder_hidden_size * self.encoder.num_directions
        self.context2decoder_h = layers.FeedForward(context_input_size, config.decoder_hidden_size,
                                                 num_layers=1, activation=config.activation)
        self.context2decoder_c = layers.FeedForward(context_input_size, config.decoder_hidden_size,
                                                 num_layers=1, activation=config.activation)
```

models/speakaddr.py

SpeakAddr - Code

```
class LSTMCell(nn.Module):
    def __init__(self, input_size, hidden_size):
        super(LSTMCell, self).__init__()
        self.input_size = input_size
        self.hidden_size = hidden_size

        self.u12u = nn.Linear(input_size, input_size, bias=False)
        self.u22u = nn.Linear(input_size, input_size, bias=False)
        self.x2h = nn.Linear(input_size, 4 * hidden_size, bias=True)
        self.u2h = nn.Linear(input_size, 4 * hidden_size, bias=True)
        self.h2h = nn.Linear(hidden_size, 4 * hidden_size, bias=True)

    def forward(self, word_input, user1_input, user2_input, hidden):
        hx, cx = hidden

        user_gate = torch.tanh(self.u12u(user1_input) + self.u22u(user2_input))
        gates = self.x2h(word_input) + self.h2h(cx) + self.u2h(user_gate)
        ingate, forgetgate, cellgate, outgate = gates.chunk(4, 1)
        ingate = torch.sigmoid(ingate)
        forgetgate = torch.sigmoid(forgetgate)
        cellgate = torch.tanh(cellgate)
        outgate = torch.sigmoid(outgate)

        cy = torch.mul(cx, forgetgate) + torch.mul(ingate, cellgate)
        hy = torch.mul(outgate, torch.tanh(cy))

        return hy, cy
```

models/speakaddr.py

Train

Environment

- SW
 - Python 3.7.3
 - PyTorch 1.1.0
- HW
 - Intel Xeon CPU E5-2640
 - 256GB
 - GTX 1080TI

Train

Operation

```
bash RunTrainTC.sh 0 SpeakAddr 30 50 True True
```

- 0: GPU ID
- HRED: model name
- 30: batch size
- 50: maximum epoch
- True: use pretrained word vector
- True: take user data

```
export CUDA_VISIBLE_DEVICES=$1
```

```
python train.py --data=tc_10_10 --model=$2 --batch_size=$3 --eval_batch_size=$3 --n_epoch=$4 --pretrained_wv=$5 --users=$6
```

Train - Code

```
def train(self):
    epoch_loss_history = list()
    min_validation_loss = sys.float_info.max
    patience_cnt = self.config.patience

    for epoch_i in range(self.epoch_i, self.config.n_epoch):
        self.epoch_i = epoch_i
        batch_loss_history = list()
        self.model.train()
        n_total_words = 0
        for batch_i, (conversations, convs_length, utterances_length, users) in \
            enumerate(tqdm(self.train_data_loader, ncols=80)):
```

solvers/speakaddr_solver.py

Train - Code

```
input_utterances = to_var(torch.LongTensor(input_utterances))
target_utterances = to_var(torch.LongTensor(target_utterances))
input_utterance_length = to_var(torch.LongTensor(input_utterance_length))
target_utterance_length = to_var(torch.LongTensor(target_utterance_length))
  


```
self.optimizer.zero_grad()

utterances_logits = self.model(input_utterances, conv_users, input_utterance_length,
 input_conversation_length, target_utterances, decode=False)

batch_loss, n_words = masked_cross_entropy(utterances_logits, target_utterances, target_utterance_length)
```


```

solvers/speakaddr_solver.py

Train - Code

```
[nosyu@yogurt src]$ bash Run_train_TC.sh 1 SpeakAddr 30 50 True True
Vocabulary size: 14318
User size: 602
Load the wv Done
/home/nosyu/anaconda3/lib/python3.7/site-packages/torch/nn/modules/rnn.py:46: UserWarning: dropout option
, so non-zero dropout expects num_layers greater than 1, but got dropout=0.2 and num_layers=1
    "num_layers={}".format(dropout, num_layers))
Parameter initiailization
    encoder.rnn.weight_hh_l0
    encoder.rnn.bias_hh_l0
    encoder.rnn.weight_hh_l0_reverse
    encoder.rnn.bias_hh_l0_reverse
Epoch: 1, iter 0: loss = 9.578
Epoch: 1, iter 100: loss = 6.180
Epoch: 1, iter 200: loss = 6.018
Epoch: 1, iter 300: loss = 6.135
Epoch: 1, iter 400: loss = 6.058
Epoch: 1, iter 500: loss = 5.963
Epoch: 1, iter 600: loss = 5.735
100%|                                         | 659/659 [01:33<00:00,  7.25it/s]
Epoch 1 loss average: 6.310
Save parameters to /data/nosyu/git/conversation-models/results/tc_10_10/SpeakAddr/20190729_213302/1.pkl

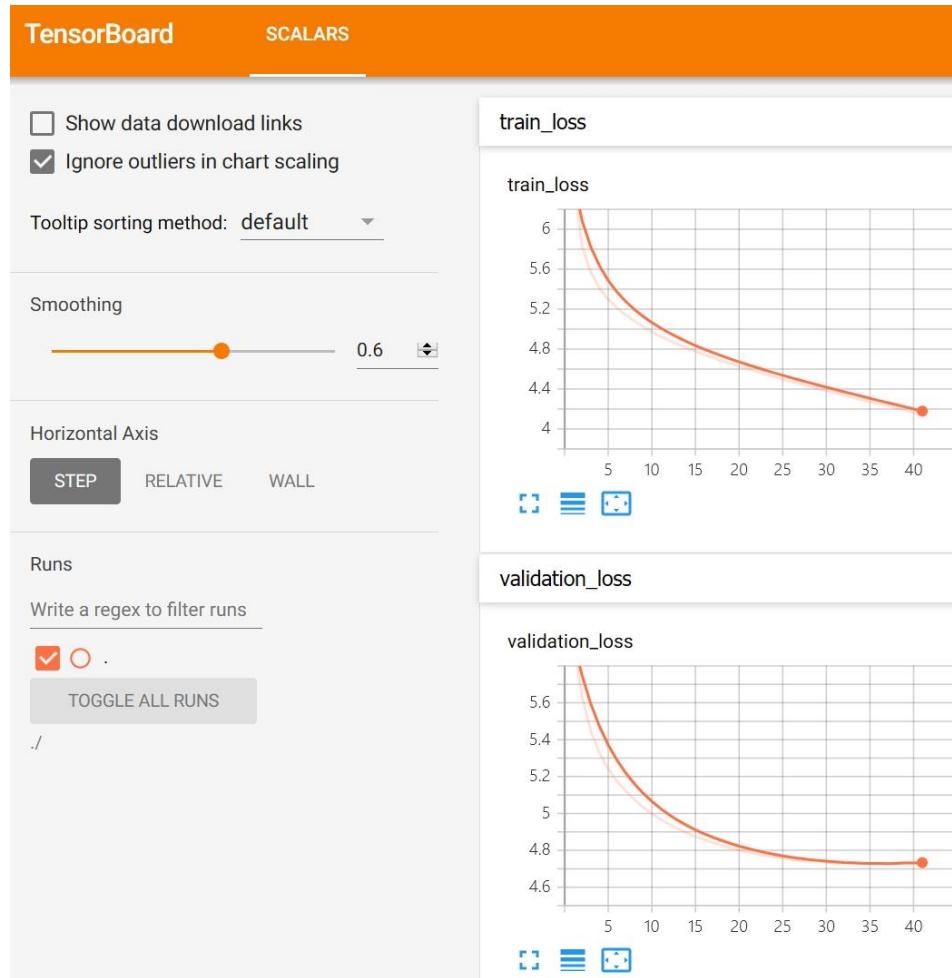
<Validation>...
100%|                                         | 81/81 [00:03<00:00, 23.50it/s]
Validation loss: 5.804

Epoch: 2, iter 0: loss = 5.809
Epoch: 2, iter 100: loss = 5.785
Epoch: 2, iter 200: loss = 5.775
Epoch: 2, iter 300: loss = 5.687
Epoch: 2, iter 400: loss = 5.692
```

Train - Result

```
[nosyu@yogurt 20190729_155444]$ ls -la
total 785968
drwxrwxr-x 2 nosyu nosyu      301 Jul 29 19:34 .
drwxrwxr-x 4 nosyu nosyu      52 Jul 29 11:33 ..
-rw-rw-r-- 1 nosyu nosyu 72981769 Jul 29 16:07 11.pkl
-rw-rw-r-- 1 nosyu nosyu 72981769 Jul 29 15:56 1.pkl
-rw-rw-r-- 1 nosyu nosyu 72981769 Jul 29 16:19 21.pkl
-rw-rw-r-- 1 nosyu nosyu 72981769 Jul 29 16:30 31.pkl
-rw-rw-r-- 1 nosyu nosyu 72981769 Jul 29 16:36 35.pkl
-rw-rw-r-- 1 nosyu nosyu 72981769 Jul 29 16:37 36.pkl
-rw-rw-r-- 1 nosyu nosyu 72981769 Jul 29 16:38 37.pkl
-rw-rw-r-- 1 nosyu nosyu 72981769 Jul 29 16:39 38.pkl
-rw-rw-r-- 1 nosyu nosyu 72981769 Jul 29 16:40 39.pkl
-rw-rw-r-- 1 nosyu nosyu 72981769 Jul 29 16:41 40.pkl
-rw-rw-r-- 1 nosyu nosyu 72981769 Jul 29 16:41 41.pkl
-rw-rw-r-- 1 nosyu nosyu     2194 Jul 29 15:54 60...19.txt
-rw-rw-r-- 1 nosyu nosyu     4166 Jul 29 16:41 events.out.tfevents.1564383295.yogurt.uilab.kr
-rw-rw-r-- 1 nosyu nosyu    994373 Jul 29 19:34 responses_test_1_1_5_35.txt
-rw-rw-r-- 1 nosyu nosyu   1014278 Jul 29 17:45 responses_test_1_1_5_37.txt
```

Train - Result



Inference

Operation

```
bash RunExportTestSamples.sh 0 SpeakAddr 30 True True 1 1 30.pkl
```

- 0: GPU ID
- SpeakAddr: model name
- 30: batch size
- True: use pretrained word vector
- True: take user data
- 1: number of context utterance
- 1: number of sample step
- 30.pkl: saved model file

Inference - Result

```
[nosyu@yogurt 20190729_155444]$ ls -la
total 785968
drwxrwxr-x 2 nosyu nosyu      301 Jul 29 19:34 .
drwxrwxr-x 4 nosyu nosyu      52 Jul 29 21:33 ..
-rw-rw-r-- 1 nosyu nosyu 72981769 Jul 29 16:07 11.pkl
-rw-rw-r-- 1 nosyu nosyu 72981769 Jul 29 15:56 1.pkl
-rw-rw-r-- 1 nosyu nosyu 72981769 Jul 29 16:19 21.pkl
-rw-rw-r-- 1 nosyu nosyu 72981769 Jul 29 16:30 31.pkl
-rw-rw-r-- 1 nosyu nosyu 72981769 Jul 29 16:36 35.pkl
-rw-rw-r-- 1 nosyu nosyu 72981769 Jul 29 16:37 36.pkl
-rw-rw-r-- 1 nosyu nosyu 72981769 Jul 29 16:38 37.pkl
-rw-rw-r-- 1 nosyu nosyu 72981769 Jul 29 16:39 38.pkl
-rw-rw-r-- 1 nosyu nosyu 72981769 Jul 29 16:40 39.pkl
-rw-rw-r-- 1 nosyu nosyu 72981769 Jul 29 16:41 40.pkl
-rw-rw-r-- 1 nosyu nosyu 72981769 Jul 29 16:41 41.pkl
-rw-rw-r-- 1 nosyu nosyu     2194 Jul 29 15:54 config.txt
-rw-rw-r-- 1 nosyu nosyu     4166 Jul 29 16:41 events.out.events.1504383295.yogurt.ulab.kr
-rw-rw-r-- 1 nosyu nosyu    994373 Jul 29 19:34 responses_test_1_1_5_35.txt
-rw-rw-r-- 1 nosyu nosyu   1014278 Jul 29 17:45 responses_test_1_1_5_37.txt
```

Inference - Result



if i watch this ncoe vid any
more i'm going to go blind
he is so hot in this

me too !
i love it too !



me too !
me too !



Inference - Result



hi how are you feeling ?
i was so worried about you

awww , i am so
glad to hear it .



awww , i am so
glad to see you .



Conclusion

- Described the conversation corpus
- Introduced conversation models with code
 - HRED
 - SpeakAddr
- Showed the way to run the code and results

Thank you!
Any questions or comments?

JinYeong Bak
jy.bak@kaist.ac.kr
U&I Lab, KAIST