

Unsupervised Counselor Dialogue Clustering for Positive Emotion Elicitation in Neural Dialogue System

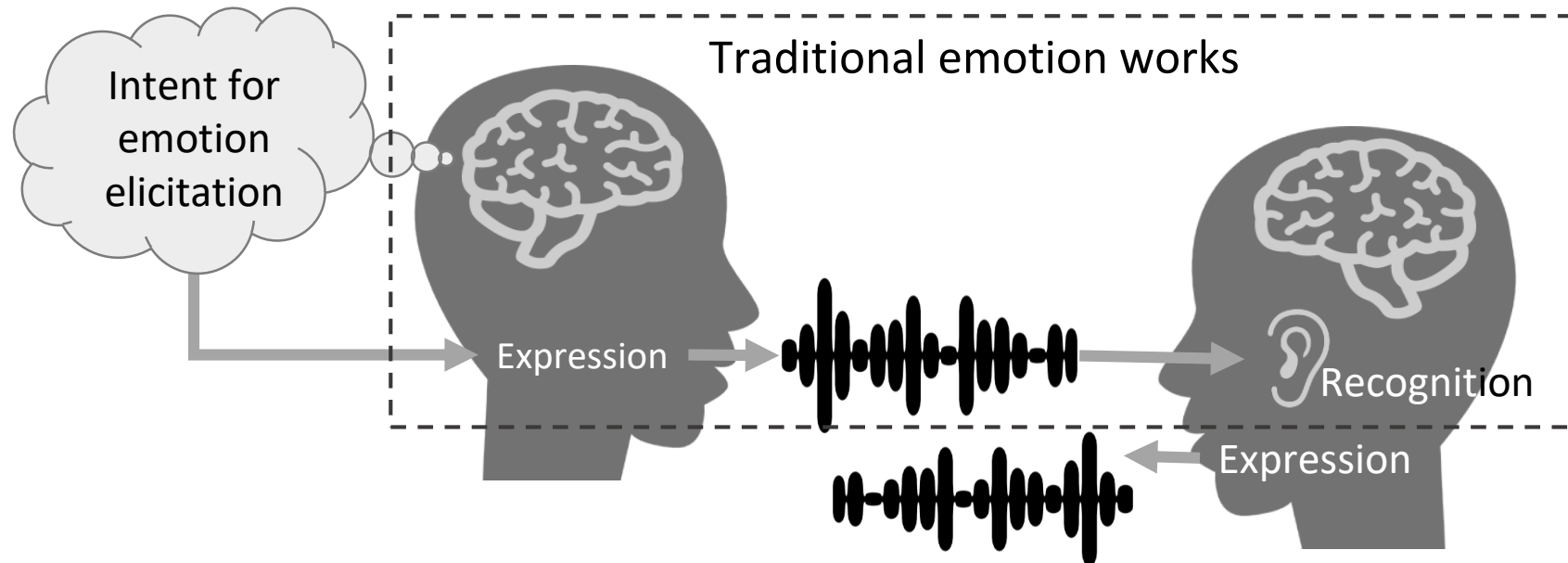
Nurul Lubis, Sakriani Sakti, Koichiro Yoshino, Satoshi Nakamura
Information Science Division, Nara Institute of Science and Technology, JAPAN

Affective dialogue systems

High potential of dialogue systems to address the emotional needs of users

- Increase of dialogue system works and applications in various tasks involving affect
 - Companion for the elderly
[Miehle et al., 2017]
 - Distress clues assessment
[DeVault et al., 2014]
 - Affect-sensitive tutoring
[Forbes-Riley and Litman, 2012]

Emotion elicitation



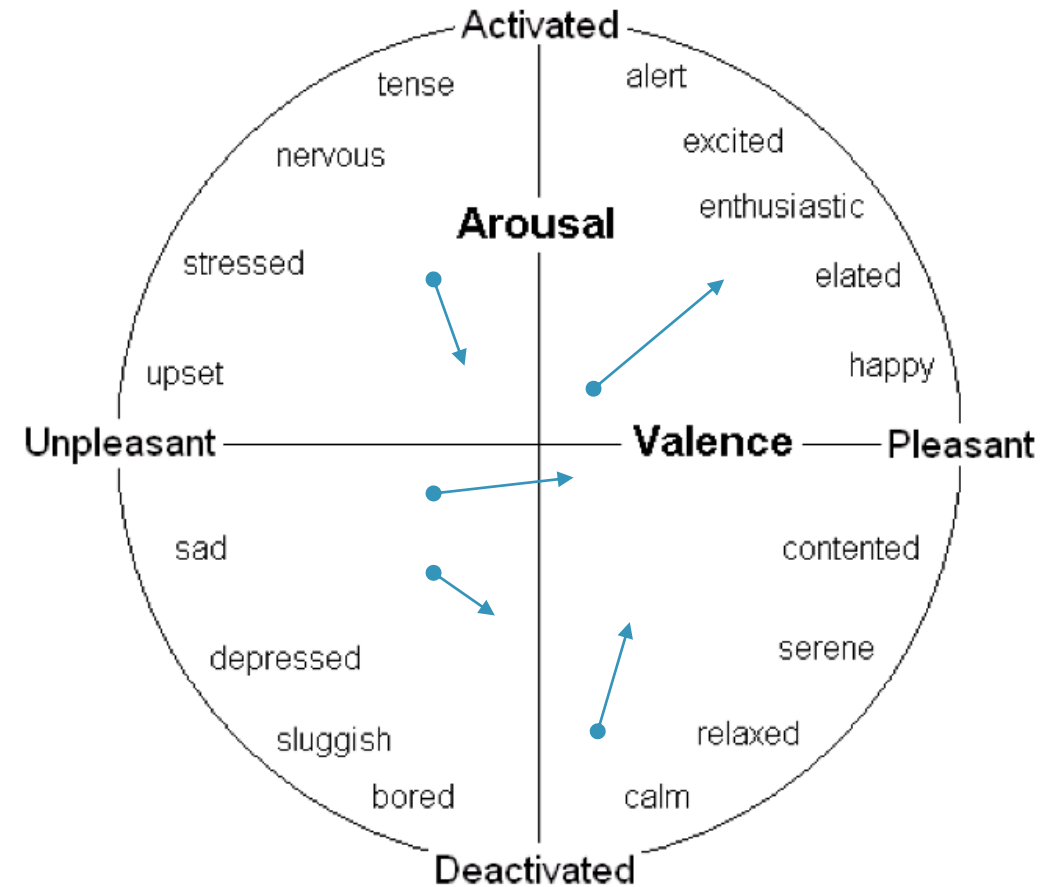
Emotion elicitation: eliciting change of emotion in dialogue

- Using machine translation with target emotion (Hasegawa et al., 2013)
- Using system's affective personalities (Skowron et al., 2013)
- ✗ Have not yet considered the **emotional benefit** for the user

Positive emotion elicitation

We aim to draw on an overlooked potential of emotion elicitation to **improve user emotional states**

- A chat-based dialogue system with an implicit goal of **positive emotion elicitation**



Circumplex model of affect [Russell, 1980]

Different responses elicit different emotions



I failed the test.

I failed the test.

Oh, again?

Yeah...

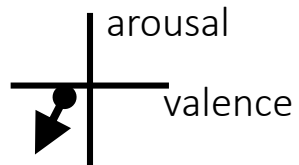
I failed the test.

You will do better next time!

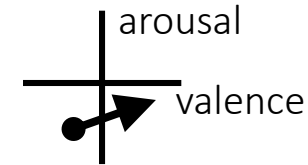
Thank you.

Emotional impact

Negative



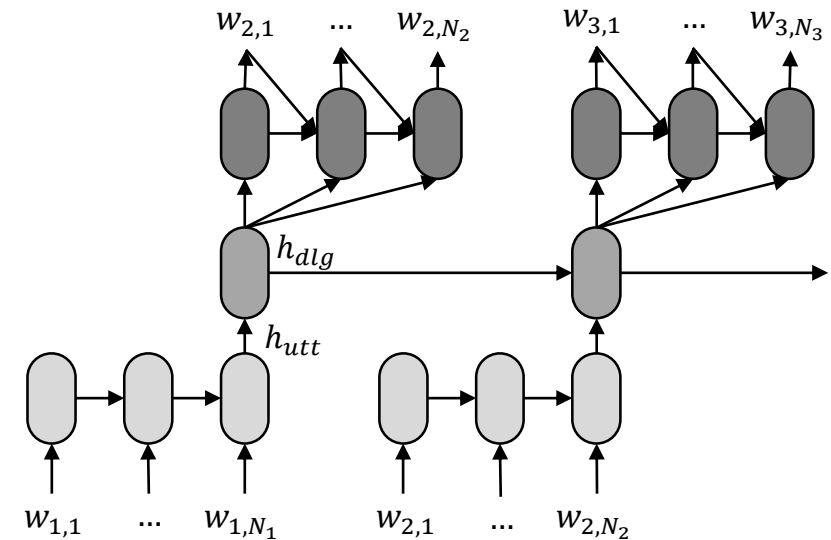
Positive



Neural chat-based dialogue system

- RNN encoder-decoder [Vinyals et al., 2015]
- Hierarchical recurrent encoder-decoder (HRED) [Serban et al., 2016]
- Generating dialogue response with emotional expression [Zhou et al., 2018]

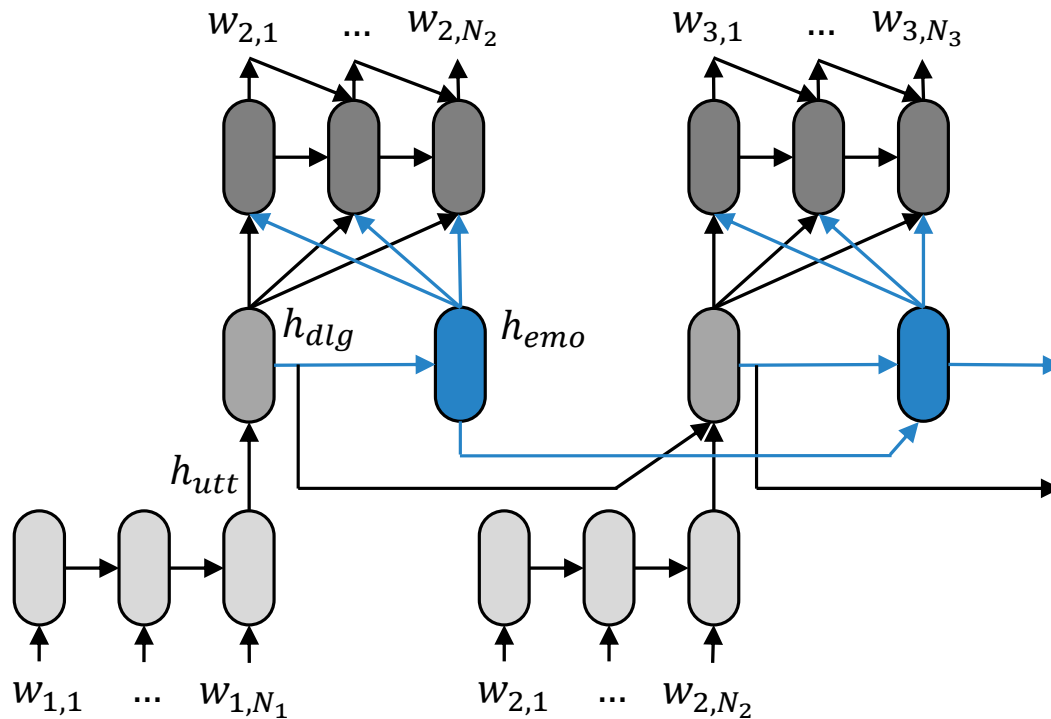
Application towards emotion elicitation is still very lacking.



[Serban et al., 2016]

Emotion-sensitive response generation: Emo-HRED

[Lubis et al., 2018] in Proc. AAAI 2018



- Encodes emotional context and considers it in generating a response
- Training data contains responses that elicit positive emotion
- Significant improvement on perceived emotional impact

Limitations

1. Has not yet learned strategies from an expert
2. Short and generic responses with positive-affect words

Challenge and proposal

1. Goal: Learning elicitation strategy from an expert

- Challenge: Absence of data that shows
 - positive emotion elicitation in everyday situations
 - expert strategy in affective dialogue
- Proposed: Construct a dialogue **corpus** involving an **expert** in a positive emotion elicitation scenario

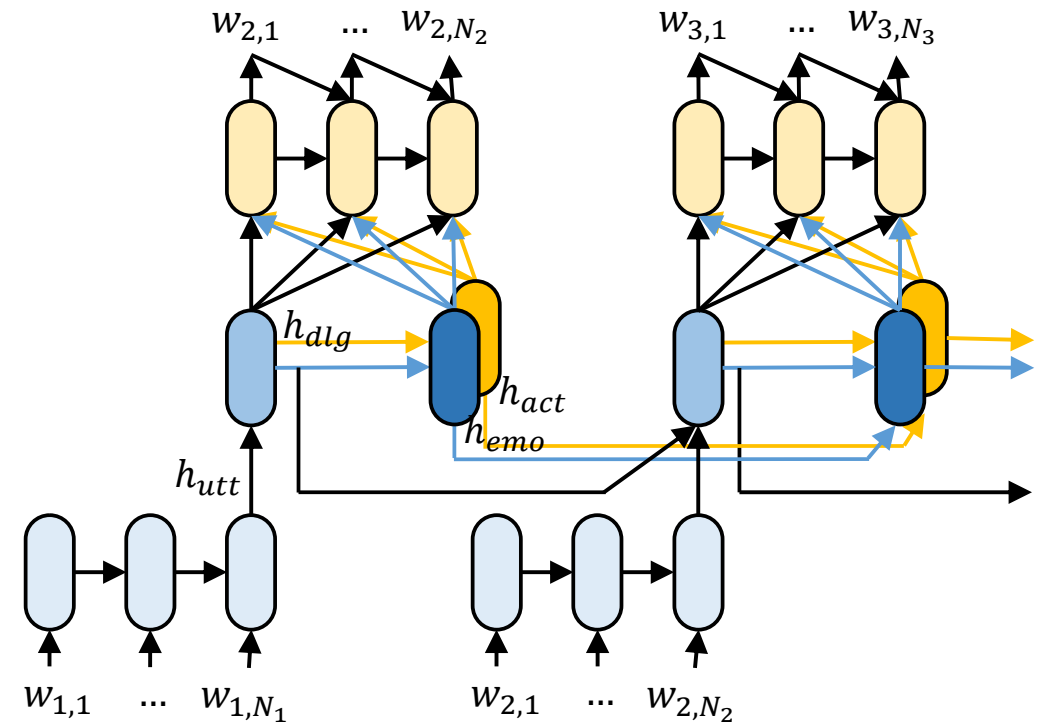
2. Goal: increase variety in the generated response to improve engagement

- Challenge: Data sparsity
- We hypothesize that higher level information, e.g. dialogue action, will reduce data sparsity
 - categorizing responses
 - emphasizing this information in the training and generation process.

Proposed architecture: Multi context HRED (MC-HRED)

A neural dialogue system which generate response based on multiple dialogue contexts

- Dialogue history
- User emotional state
- Response action label



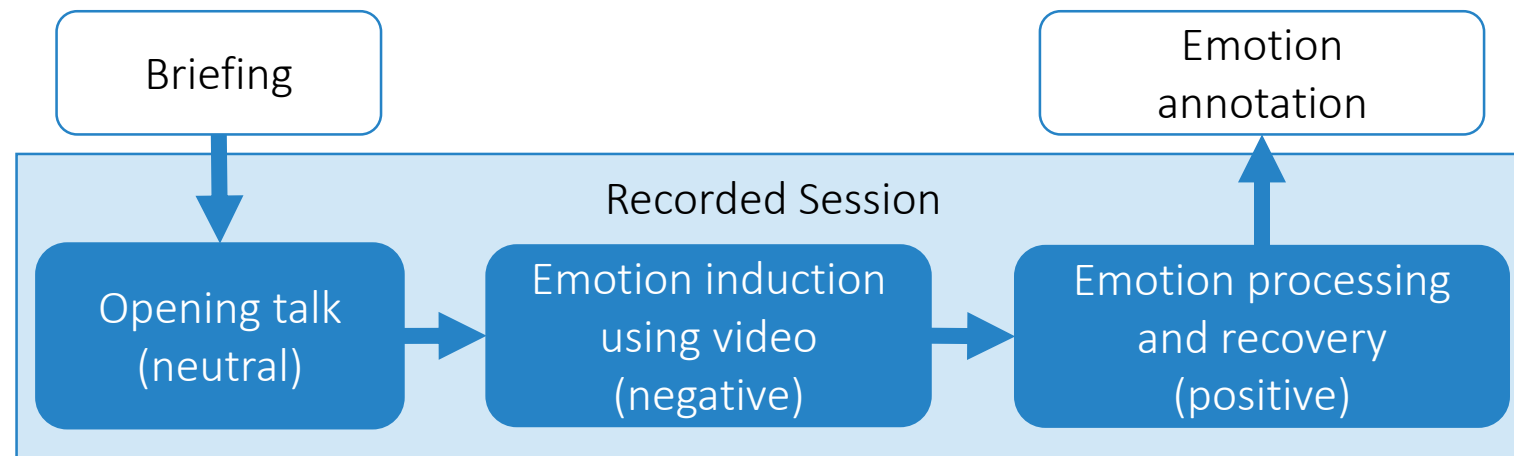
MC-HRED architecture.

Corpus construction

Positive Emotion Elicitation by an Expert

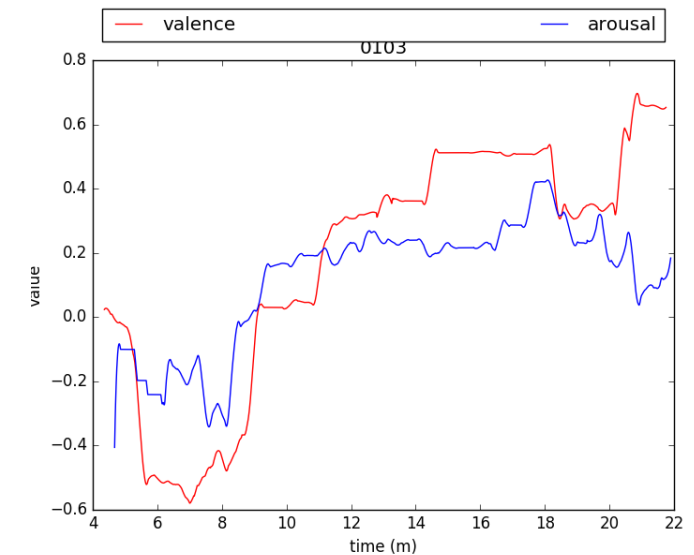
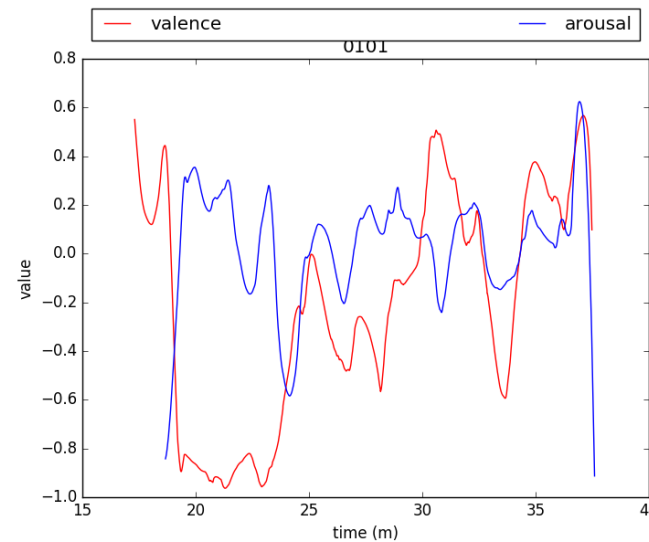
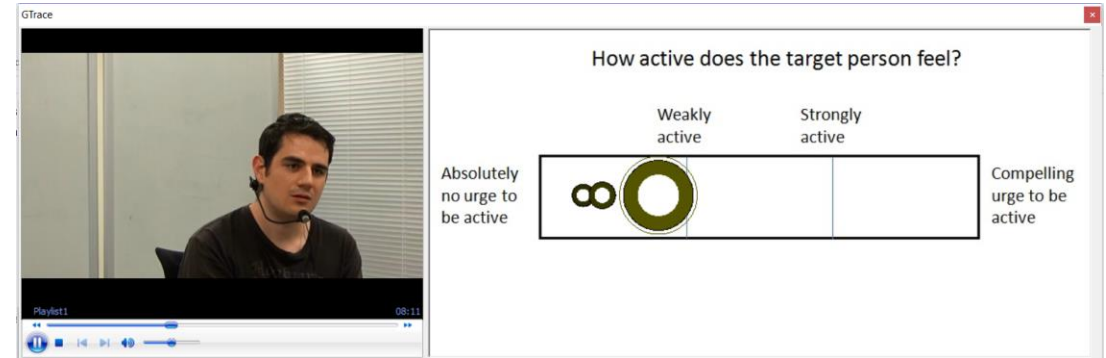
Data recording design

- Goal: learn expert strategy for eliciting positive emotion
- Collect:
 - Interaction between an **expert** and a **participant**
 - Condition the interaction with negative emotion
 - Expert guides the conversation to allow participant's emotion recovery and reinstate positive emotion



Data collection and annotation

- 60 sessions: 23 hours and 41 minutes of material
 - 1 counselor, 30 participants
 - 2 sessions per participant
 - 1 induced to anger
 - 1 induced to sadness
- Self-report emotion annotation using Gtrace [Cowie et al., 2000]
- Transcription



Unsupervised Clustering of Counselor Dialogue

Counselor dialogue clustering

Goal: To find high-level information

- Information equivalent to dialogue acts
- Specific to the dialogue scenario
- Retaining affective intents

✗ Human annotation

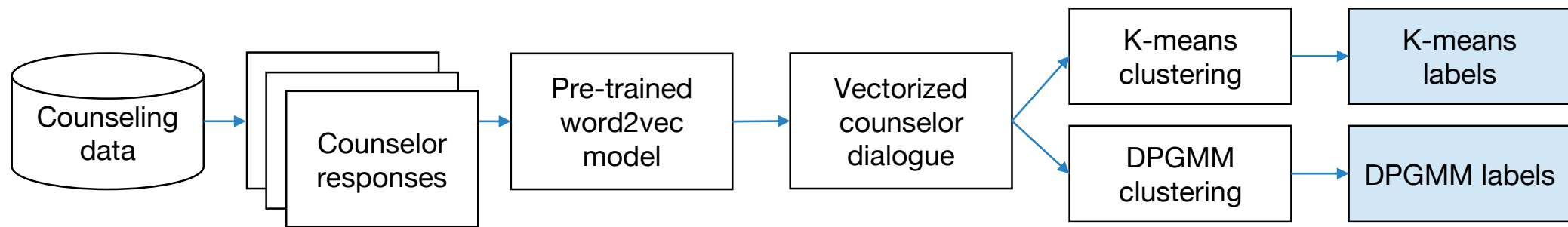
- Expensive, labor intensive
- Low reliability

✗ Standard dialogue acts classifier

- May not cover specific emotion-related intent in the data

✓ Unsupervised clustering

Counselor dialogue clustering



K-Means

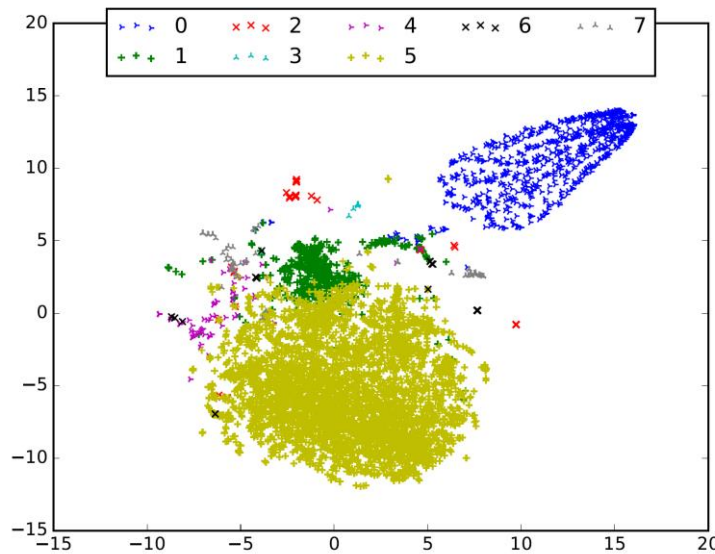
- Need to predefine how many clusters
- We choose K empirically

DPGMM

- No prior definition of model complexity

Analysis

K-means: 8 clusters



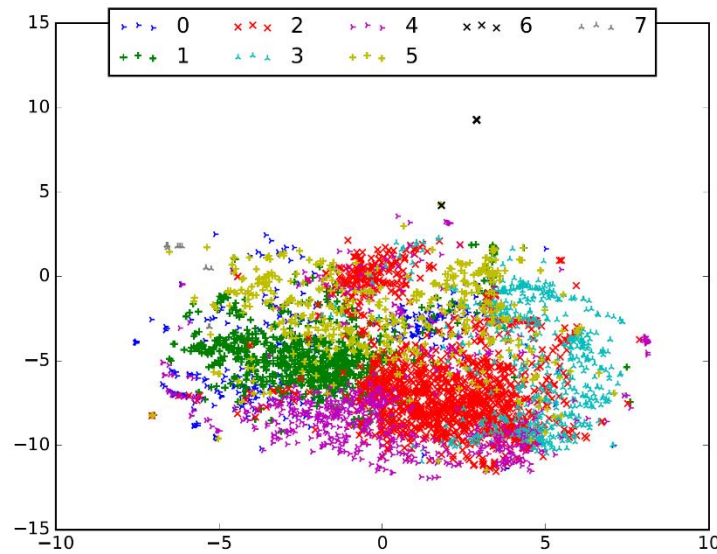
“So you feel frustrated.”
“I guess we all have to be careful.”

“Maybe, yes yes.”
“Right.”

“Mm mm.”
“Yes, hm mm”

13 July 2018

K-means: 8 sub-clusters

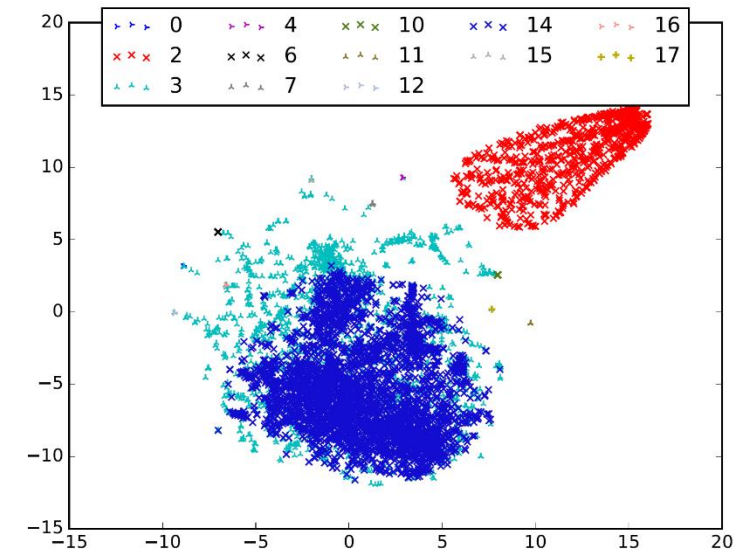


“Have you thought about this
kind of issue before?”

“So who do you think is
responsible for it?”

Nurul Lubis

DPGMM: 13 clusters



An assortment of
shorter sentences

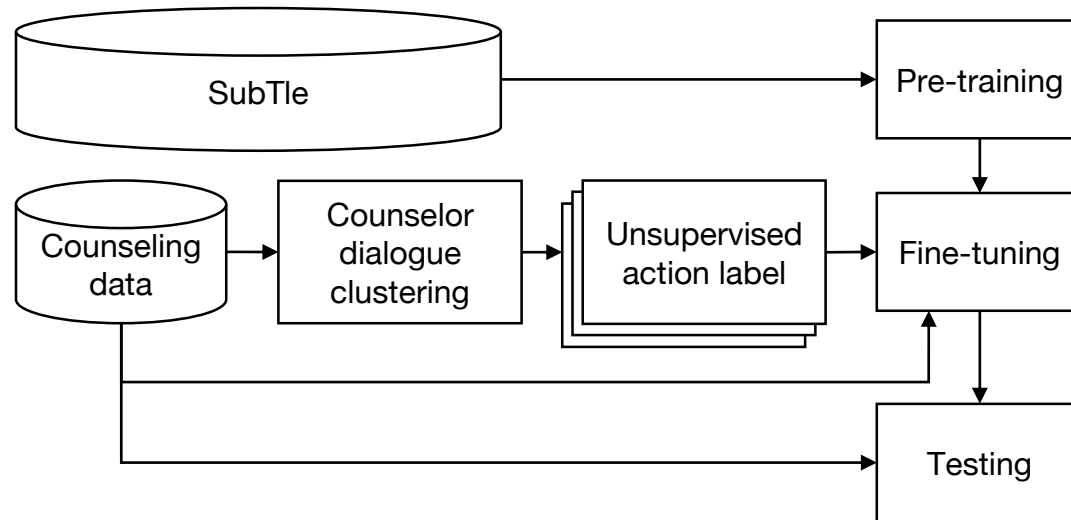
An assortment of
longer sentences

“Mm mm.”
“Yes, hm mm”

16

Experiment

Experimental set-up



Pre-training

- SubTle corpus [Ameixa et al., 2014] ~5.5M dialogue pairs from movie subtitle
- HRED to retain information across dialogue turns

Fine-tuning

- Counseling corpus
- Baseline: Emo-HRED
 - emotion context
- Proposed: MC-HRED
 - emotion and action contexts
- Clust-HRED
 - Action context

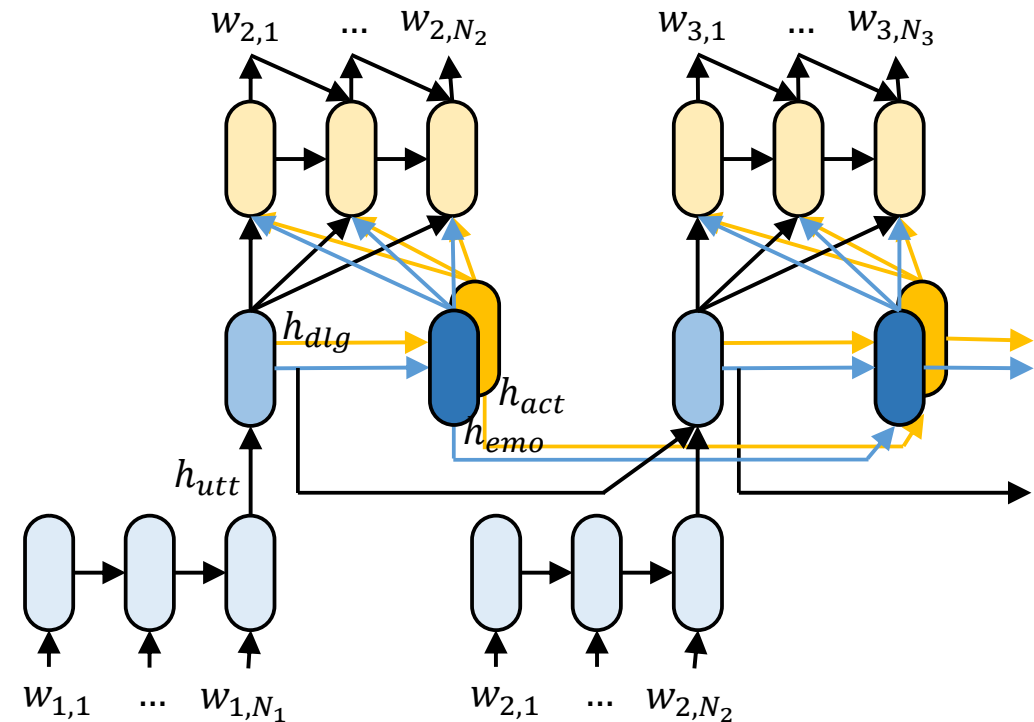
Pre-training and fine-tuning

Pre-training initializes the weights of HRED components

Selective fine-tuning: only optimize parameters affected by new contexts

MC-HRED is jointly trained on combined losses

- NLL of target response
- Emotion prediction error
- Action prediction error



MC-HRED architecture.

Objective evaluation: perplexity

Model	Emo	Action	Perplexity
Emo-HRED	yes	no	42.60
Clust-HRED	no	K-means	39.57
		DPGMM	30.57
MC-HRED	yes	K-means	29.57
		DPGMM	32.04

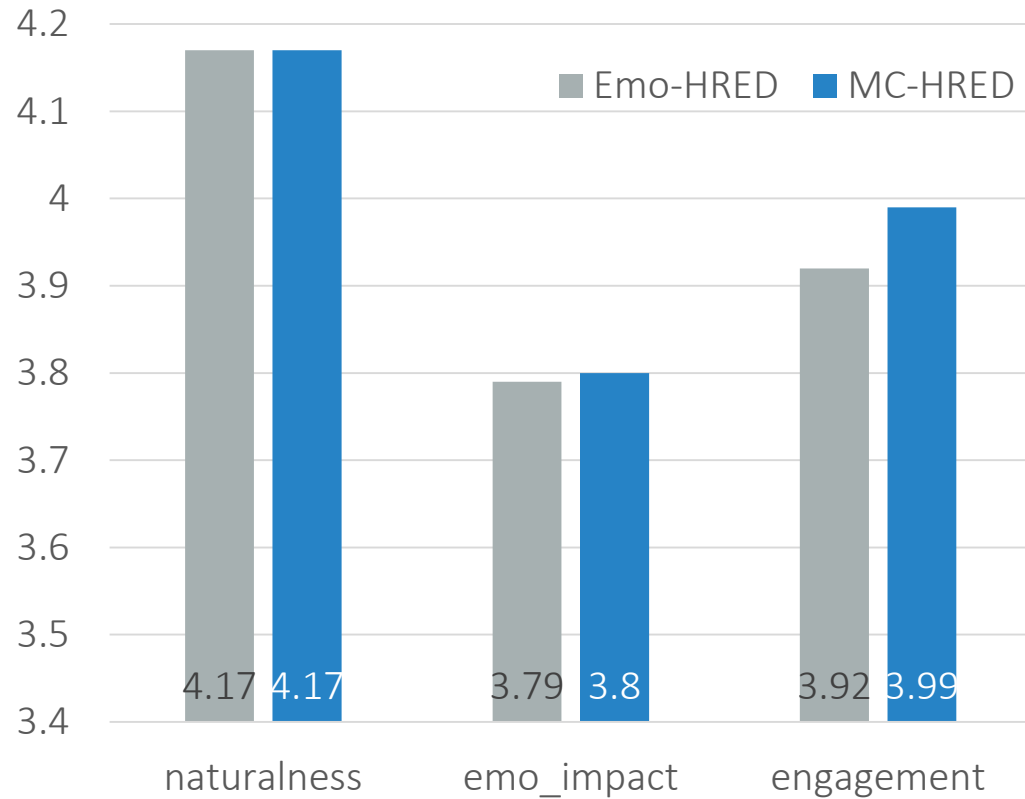
- Combining cluster label and emotion contexts
 - K-means cluster label shows improvements
 - DPGMM cluster label slightly worsen

Objective evaluation: perplexity

Model	Emo	Action	Perplexity		
			all	short	long
Emo-HRED	yes	no	42.60	35.74	61.17
Clust-HRED	no	K-means	39.57	32.30	57.37
		DPGMM	30.57	24.79	42.25
MC-HRED	yes	K-means	29.57	23.23	38.73
		DPGMM	32.04	25.00	42.43

- Perplexity on short and long queries
- Performance on short queries are consistently better than long ones
- MC-HRED with K-means obtains substantial improvement on long triples
- The multiple contexts help, especially for long inputs

Subjective evaluation



- 100 queries, each judged by 20 crowd workers
 - Naturalness
 - Emotional impact
 - Engagement
- Improved engagement while maintaining the emotional impact and naturalness
- MC-HRED produce responses with 2.53 more words on average

Conclusion

We presented

- A corpus showing expert strategy in positive emotion elicitation
- Unsupervised clustering of expert dialogue
- A multi-context neural response generation
 - Improves performance on longer queries
 - Improves dialogue engagement
 - Produces longer responses

Future Work

- Multimodal information: speech, visual
- Evaluation through user interaction

Thank you!

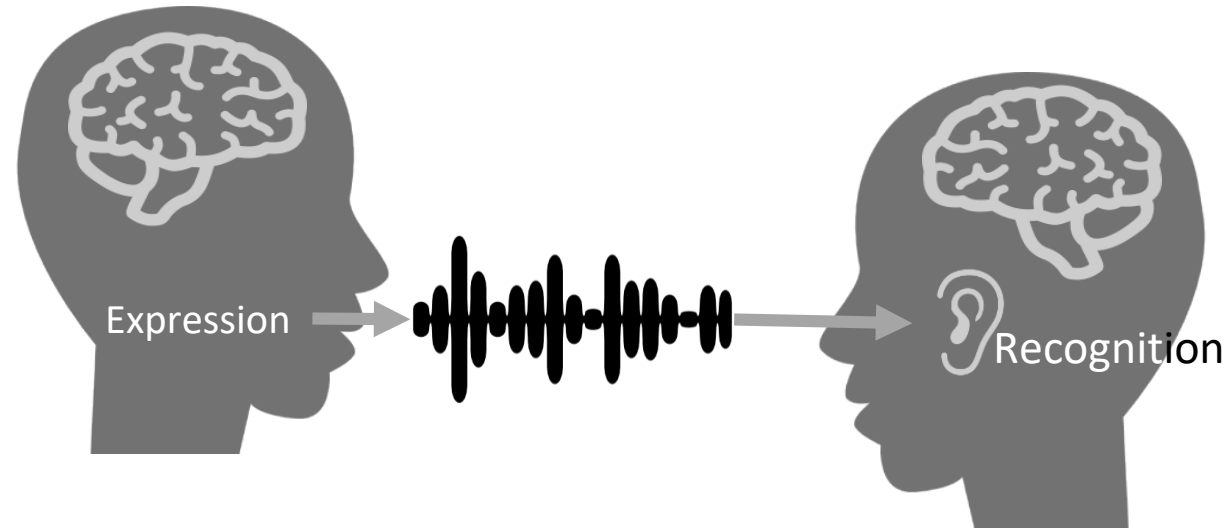
Examples

U ₁	oh how do you feel about that one.
U ₂	yes i heard the story.
U ₃ (Target)	you heard it before.
Emo-HRED	right.
MC-HRED	it's a big thing.

U ₁	are you a student here?
U ₂	uh yes, actually I just got, er that's my lab over there in social computing yes (laughter).
U ₃ (Target)	oh really. so you've been watching us going by.
Emo-HRED	Oh okay.
MC-HRED	(laughter) it's nice to meet you.

Traditional Works on Emotion

Expression and Recognition



Emotion expression or simulation

- Conveying emotion to user
- Increasing closeness and satisfaction [Higashinaka et al., 2008]

Emotion recognition

- Recognizing user's emotional state
- Increasing task success [Forbes-Riley and Litman, 2012]