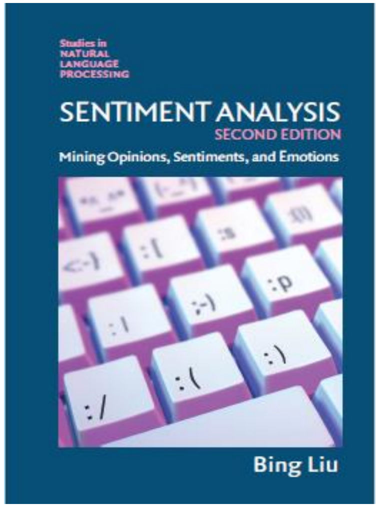
Bot Emotion

2nd edition



Bing Liu
Department of Computer Science
University of Illinois at Chicago

Introduction

- Humans are affective or emotional beings
  - feelings and emotions are an integral part of us.
- Emotion intelligence: a vital aspect of human intelligence,
  - the ability to perceive, understand, integrate, and regulate emotions
- Goal: **to create chatbots that can perceive and express emotions** and learn continually during conversations.
  - I got interested in **emotional bots** in 2015–2016 due to
    - my long interest in sentiment/emotion analysis, and
    - the fact that chatbots started to become popular at that time.

Outline

- Emotion and Definition
- Emotion Expressions
- Chatting Needs Emotions – HCI Research
- Emotional Chatting Machine: A Deep Learning Model
- Summary

Disagreements on basic emotions from theorists

Source	**Basic Emotions**
(Arnold, 1960)	Anger, aversion, courage, dejection, desire, despair, fear, hate, hope, love, sadness,
(Ekman *et al.*, 1982)	Anger, disgust, fear, joy, sadness, surprise
(Gray, 1982)	anxiety, joy, rage, terror,
(Izard, 1971)	Anger, contempt, disgust, distress, fear guilt, interest, joy, shame, surprise
(James, 1884)	Fear, grief, love, rage
(McDougall, 1926)	Anger, disgust, elation, fear, subjection, tender-emotion, wonder
(Mowrer, 1960)	Pain, pleasure
(Oatley and Jobnson-Laird, 1987)	Anger, disgust, anxiety, happiness, sadness
(Panksepp, 1982)	Expectancy, fear, rage, panic
(Plutchik, 1980)	Acceptance, anger, anticipation, disgust, joy, fear, sadness, surprise

Primary, Secondary and Tertiary emotions

(Parrott, 2001)

Primary emotion	Secondary emotion	Tertiary emotion
Anger	Disgust	Contempt, Loathing, Revulsion
	Envy	Jealousy
	Exasperation	Frustration
	Irritability	Aggravation, Agitation, Annoyance, Crosspatch, Grouchy, Grumpy
	Rage	Anger, Bitter, Dislike, Ferocity, Fury, Hatred, Hostility, Outrage, Resentment, Scorn, Spite, Vengefulness, Wrath
	Torment	Torment
Fear	Horror	Alarm, Fear, Fright, Horror, Hysteria, Mortification, Panic, Shock, Terror
	Nervousness	Anxiety, Apprehension (Fear), Distress, Dread, Suspense, Uneasiness, Worry

Emotion definition

(Liu 2015, 2020)

- Emotion: (e, a, m, f, c, t), - target (e, a)

  where
  - e is the target entity,
  - a is the target aspect of e that is responsible for the emotion,
  - m is the emotion type or a pair representing an emotion type and an intensity level,
  - f is the feeler of the emotion,
  - c is the cause of the emotion, and
  - t is the time when the emotion is felt or expressed.

Emotion example

- "*I am so mad with the hotel manager because he refused to refund my booking fee,*"
  - (e, a, m, f, c, t),
    - (the hotel, manager, anger, I, 'he refused to refund my booking fee', date-x)
  - Target: the hotel manager

Outline

- Emotion and Definition
- Emotion Expressions
- Chatting Needs Emotions – HCI Research
- Emotional Chatting Machine: A Deep Learning Model
- Summary

Speaker emotion

- Speaker emotion is conveyed through paralinguistic mechanisms such as
  - intonations,
  - facial expressions,
  - body movements,
  - biophysical signals or changes,
  - gestures, and posture

  - + choice of grammatical and lexical expressions

Writer emotion

- In writing, we express emotions using
  - punctuation (e.g., repeated exclamation marks, !!!!),
  - capitalization of all letters of a word,
  - emoticons,
  - lengthening of words (e.g., *slooooooow*), etc.

  - + choice of grammatical and lexical expressions

Linguistic expressions of emotion

- Emotion words/phrases: *love*, *disgusting*, *angry*, and *upset*
- Intensifiers: *very*, *so*, *extremely*, *dreadfully*, *really*, *awfully* (e.g., *awfully bad*), *on earth* (e.g., "*What on earth are you doing*?"), *the hell* (e.g., "*What the hell are you doing*?"), etc.
- Emotion-related behaviors: "*He cried after he saw his mother*"
- Superlatives: "*This car is simply the best.*" pejorative, "*He is a fascist.*", laudatory, "*He is a saint*", and sarcastic, "*What a great car, it broke the second day*".
- Swearing, cursing, insulting, blaming, accusing, etc

Cognitive gap

- Cognitive gap: people’s true psychological states of mind and the language that they use to express such states
  - may not fully match
  - Identifying the true emotion can be challenging
    - Context and multi-modal data may be needed
- Many reasons (e.g., *being polite*, and *do not want others to know one’s true feeling*)
  - Language does not always represent psychological reality.
    - For example, when one says “*I am happy with this car*,” one may not have any emotional reaction towards the car although the emotion word *happy* is used.

Outline

- Emotion and Definition
- Emotion Expressions
- Chatting Needs Emotions – HCI Research
- Emotional Chatting Machine: A Deep Learning Model
- Summary

Studies in human-computer interaction (HCI)

- Psychology argues that humans are affective beings
  - who want their emotional and social needs to be addressed (Myers, 1989).
- Prendinger et al. (2005) showed that emotion in dialogue systems can enhance user satisfaction
  - A life-like character verbally and non-verbally expressing empathy may significantly decrease user frustration and stress.
  - A life-like character with affective behavior may have a positive effect on the users' perception of task difficulty.
- Bots must detect & express emotions in chatting

Emotions → fewer breakdowns in dialogues

- Martinovski & Traum (2003) showed while bots are error-prone, this is not necessarily a major problem, since human dialogue also contains errors.
  - Main issue: big deviation from the user's conversational expectations.
  - Detecting emotions and responding empathically can inspire the human users to desire to cooperate rather than force them to adapt to the ‘stupid machine.’
- Conclusion: Detecting emotions and responding appropriately in dialogue systems lead to fewer breakdowns.

Multiple possible responses

- Given a user utterance, there can be multiple possible responses with different emotions.
  - User: Worst day ever. I arrived late because of the traffic.
  - Bot (no emotion): You are late.
  - Bot (happy): I am glad that you are like me.
  - Bot (sad): This is depressing.
  - Bot (disgust): This traffic is getting worse and worse.
  - Bot (sympathetic): I am sorry to know that you were caught in traffic.
  - …

Outline

- Emotion and Definition
- Emotion Expressions
- Chatting Needs Emotions – HCI Research
- Emotional Chatting Machine: A Deep Learning Model
- Summary

Building an emotion chatbot: challenges
(Zhou et al 2018)

- Hard to obtain large scale emotion labeled data.
  - Highly subjective in annotation
  - Different emotions overlaps and associated with each other.
    - Classification of emotion is inaccurate.
- Hard to balance emotions, grammaticality, and content in response generation.
  - We proposed an end-to-end model for generating emotional responses in an open-domain chatbot system
    - Emotional chatting machine (ECM).

Problem definition

- Given the user utterance $\boldsymbol{X} = (x_1, x_2, \cdots, x_n)$, and a emotion category *e* = {*Angry, Disgust, Happy, Like, Sad, Other*}.

- Goal: generate a response $\boldsymbol{Y} = (y_1, y_2, \cdots, y_m)$ by estimating the probability:

$$P(\boldsymbol{Y}|\boldsymbol{X}, e) = \prod_{t=1}^{m} P(y_t|y_{<t}, \boldsymbol{X}, e).$$

Emotional Chatting Machine (ECM)

- Based on the encoder-decoder framework of the general sequence-to-sequence (seq2seq) approach.
  - Implemented with GRU
- The encoder converts the post/utterance sequence $\boldsymbol{X} = (x_1, x_2, \cdots, x_n)$ to hidden representations $\boldsymbol{h} = (\boldsymbol{h}_1, \boldsymbol{h}_2, \cdots, \boldsymbol{h}_n)$, which is defined as:

$$\boldsymbol{h}_t = \mathbf{GRU}(\boldsymbol{h}_{t-1}, x_t).$$

The decoder takes as input a context vector $\boldsymbol{c}_t$ and the embedding of a previously decoded word $\boldsymbol{e}(y_{t-1})$ to update its state $\boldsymbol{s}_t$ using another GRU:

$$\boldsymbol{s}_t = \mathbf{GRU}(\boldsymbol{s}_{t-1}, [\boldsymbol{c}_t; \boldsymbol{e}(y_{t-1})]),$$

the decoder generates a token by sampling from the output probability distribution o_t

$$\begin{aligned} y_t \sim \boldsymbol{o}_t &= P(y_t \mid y_1, y_2, \cdots, y_{t-1}, \boldsymbol{c}_t), \\ &= \mathrm{softmax}(\mathbf{W}_\mathbf{o} \boldsymbol{s}_t). \end{aligned}$$

Overview of ECM

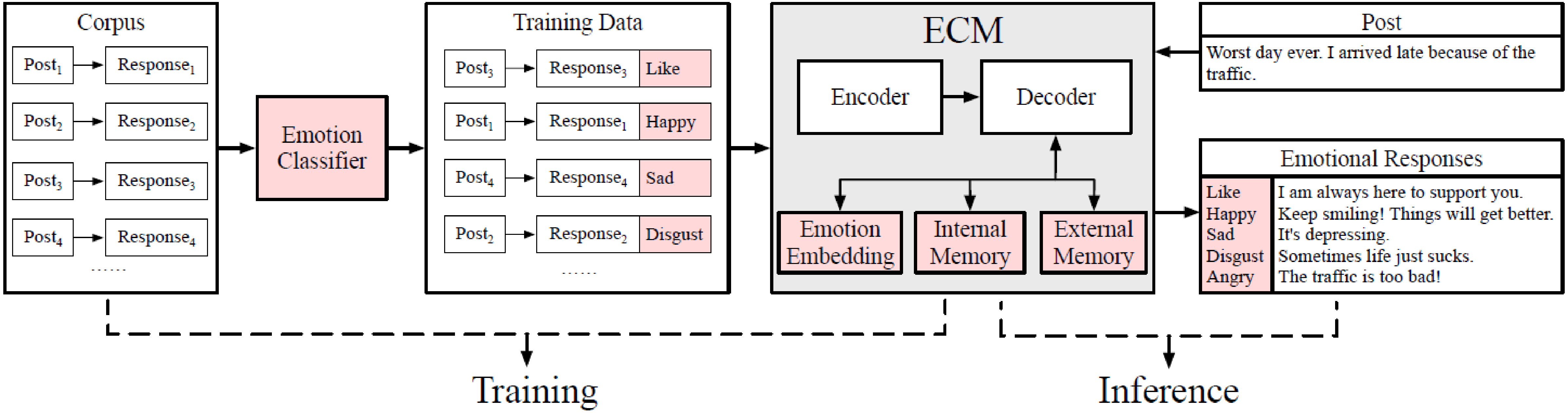


Figure 1: Overview of ECM (the grey unit). The pink units are used to model emotion factors in the framework.

Emotion category embedding

- Each emotion category e is represented by a real-valued, low dimensional vector.
- For each emotion category e, randomly initialize the vector of an emotion category v_e, and then learn the vector of the emotion category through training.

- The decoder to update the decoder's state $\boldsymbol{s_t}$:

$$s_t = \mathrm{GRU}(s_{t-1}, [c_t; e(y_{t-1}); v_e]).$$

Internal memory of emotion state

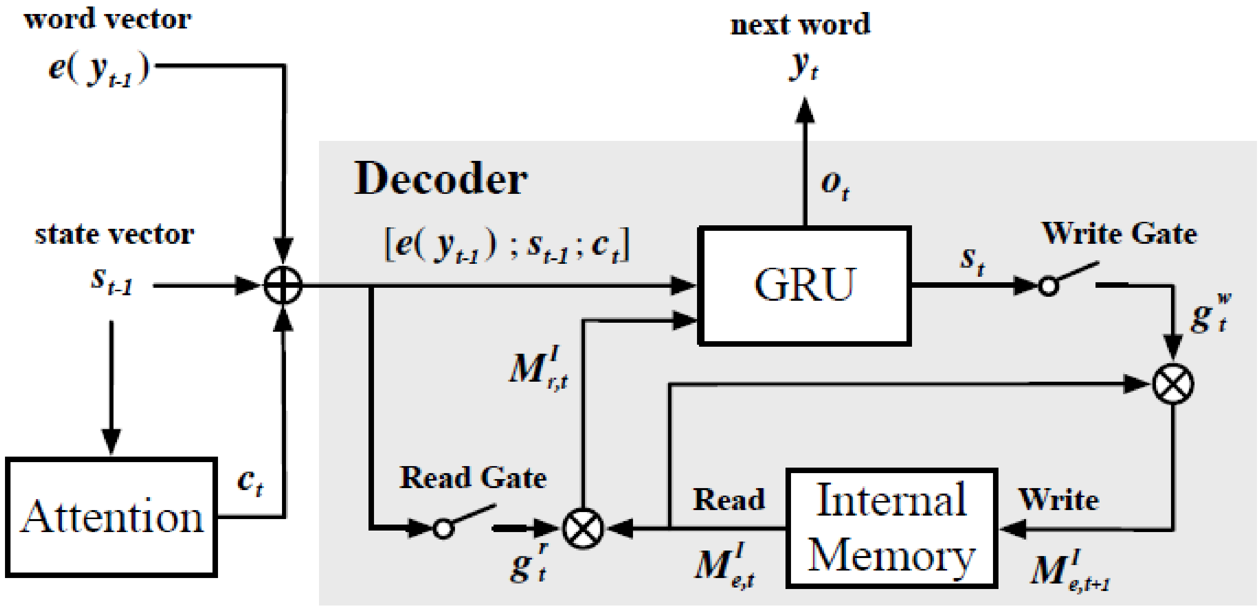


Figure 2: Data flow of the decoder with an internal memory. The internal memory $M^I_{e,t}$ is read with the read gate g^r_t by an amount $M^I_{r,t}$ to update the decoder's state, and the memory is updated to $M^I_{e,t+1}$ with the write gate g^w_t.

External memory of emotion expressions

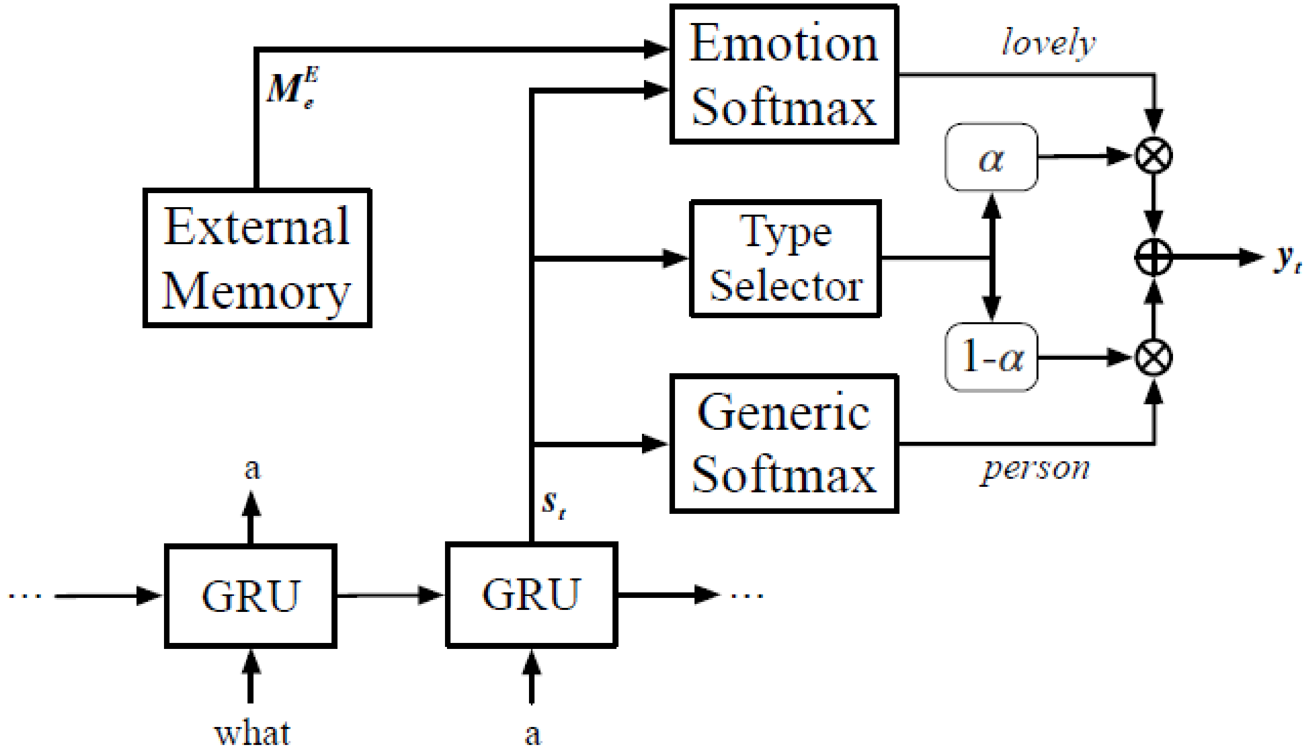


Figure 3: Data flow of the decoder with an external memory. The final decoding probability is weighted between the emotion softmax and the generic softmax, where the weight is computed by the type selector.

Evaluation

Method	Accuracy
Lexicon-based	0.432
RNN	0.564
LSTM	0.594
Bi-LSTM	0.623

- No dialogue data with emotion annotation
- An emotion annotated dataset is created.
  - An emotion classifier *E* is built using an emotion annotated dataset (NLPCC).
  - *E* automatically labels/annotates the dialogue dataset STC
- A Bi-LSTM is used to learn the emotion classifier *E*.
  - The accuracy is not high as emotion classification is challenging - annotation subjectivity and semantic overlapping of emotion categories, e.g., anger, *sadness* and *disgust*

An emotion annotated dataset ESTC

- The automatically annotated dialogue dataset is called ESTC,
  - a noisy dataset.

Training	Posts	217,905	
	Responses	Angry	234,635
		Disgust	689,295
		Happy	306,364
		Like	1,226,954
		Sad	537,028
		Other	1,365,371
Validation	Posts	1,000	
Test	Posts	1,000	

Manual evaluation

- Annotators: score each response with Content rating (0,1,2) and Emotion rating (0,1).

Method (%)	2-1	1-1	0-1	2-0	1-0	0-0
Seq2Seq	9.0	5.1	1.1	37.6	28.0	19.2
Emb	22.8	9.3	4.3	27.1	19.1	17.4
ECM	**27.2**	**10.8**	4.4	24.2	15.5	17.9

Table 5: The percentage of responses in manual evaluation with the score of *Content-Emotion*. For instance, 2-1 means content score is 2 and emotion score is 1.

Outline

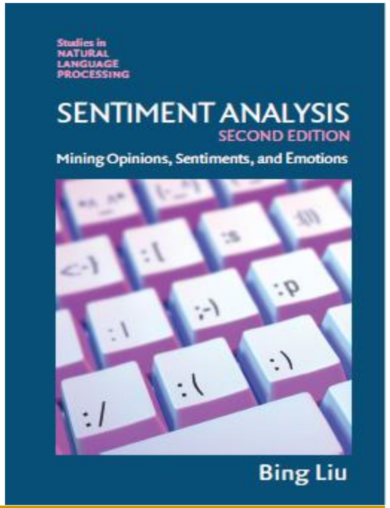
- Emotion and Definition
- Emotion Expressions
- Chatting Needs Emotions – HCI Research
- Emotional Chatting Machine: A Deep Learning Model
- Summary

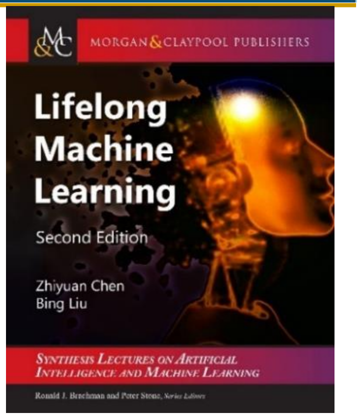
Summary

- Chatting with emotions is vital for dialogue systems.
- We discussed a deep learning model ECM (Zhou et al., 2018)
  - Quality is still not ready for prime time
  - Multi-modal emotion detection and generation
  - A bot should learn by itself continually using emotion signals (Liu, 2018)
- There are deployed bots with emotions, mainly based on rules
  - E.g., customer service and social companion bots
- This talk focused on emotions, but
  - there are also moods and other feelings (Liu, 2020)

Thank You

Q&A





- B. Liu. *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*. Cambridge University Press, 2015, 2020 (2nd edition, *forthcoming*).