

# TUA1 at the NTCIR-14 STC-3 Task

Yangyang Zhou, Zheng Liu, Xin Kang, Yunong Wu, and Fuji Ren

Tokushima University, Tokushima 770-8506, JP

{c501737062, c501837066, wuyunong}@tokushima-u.ac.jp {kang-xin, ren}@is.tokushima-u.ac.jp

## Background

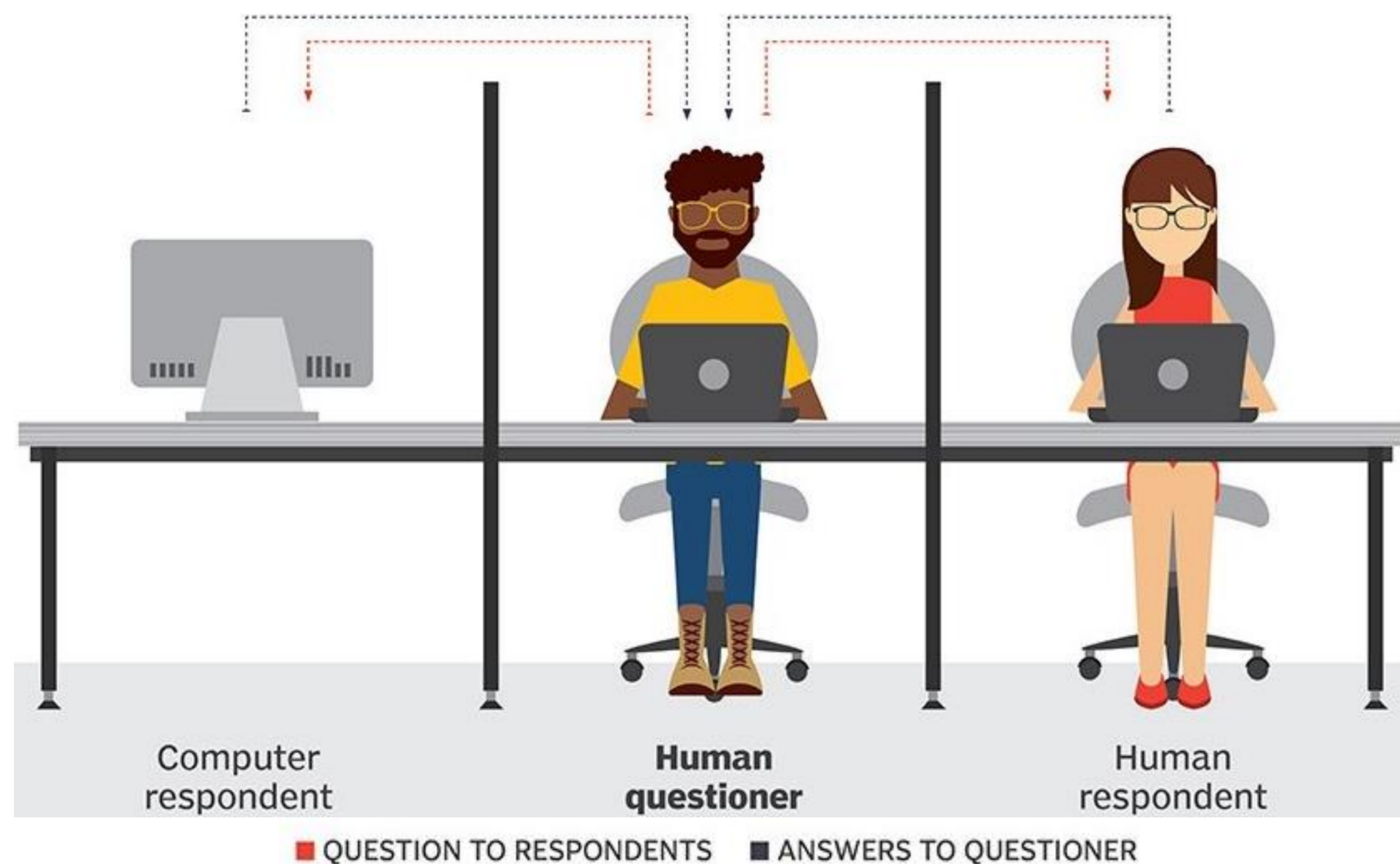


Fig.1 Turing test: If the agent wants to have human intelligence, understanding and expressing emotion is one of the keys.

Post	爱狗还会做饭的男人，最帅了! The man who cooks and loves dogs is very handsome!			
Response	Emotion Class	Coherence and Fluency	Emotion Consistency	Label
会做饭的男人是很帅的啊。 The man who cooks is handsome.	喜欢 Like	Yes	Yes	2
哈哈，我也觉得。 Haha, I feel the same way.	喜欢 Like	Yes	No	1
这是哪部电影里的？ Which movie is this from?	厌恶 Disgust	No	Yes	0
哈哈，你也是。 Haha, the same to you.	喜欢 Like	No	No	0

Table.1 The purpose of the task is to generate coherent, fluent and emotional responses based on the given posts. The first response will appear in the model parts as an example.

## P&E2R model

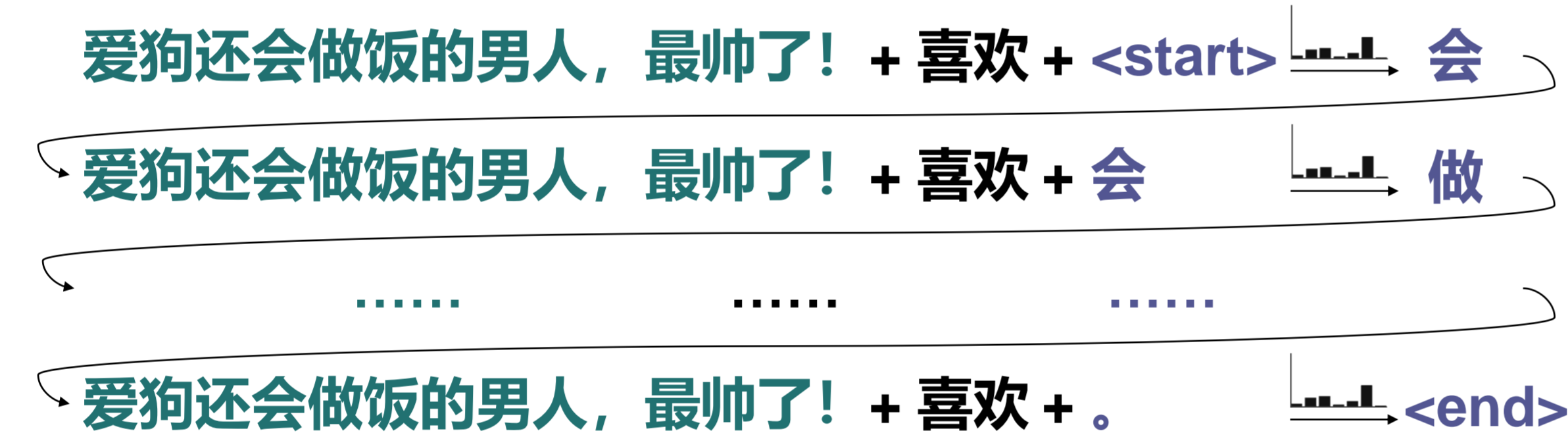
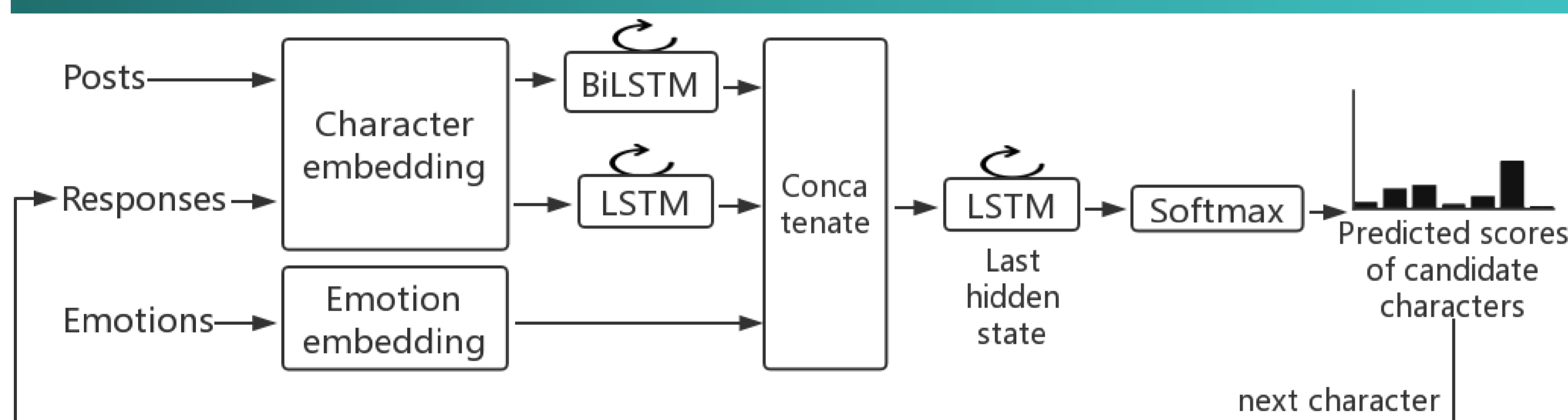


Fig.2 Structure and generation process of P&E2R model

## P2R&E2R model

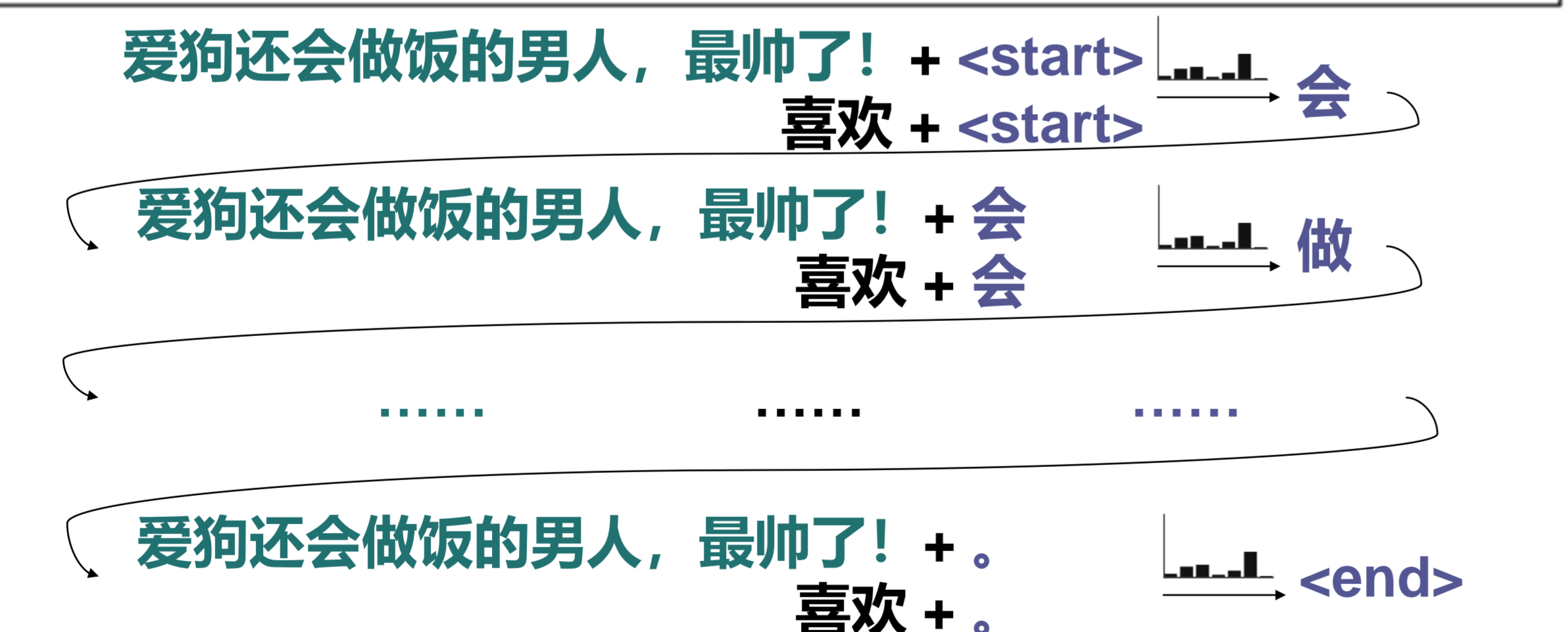
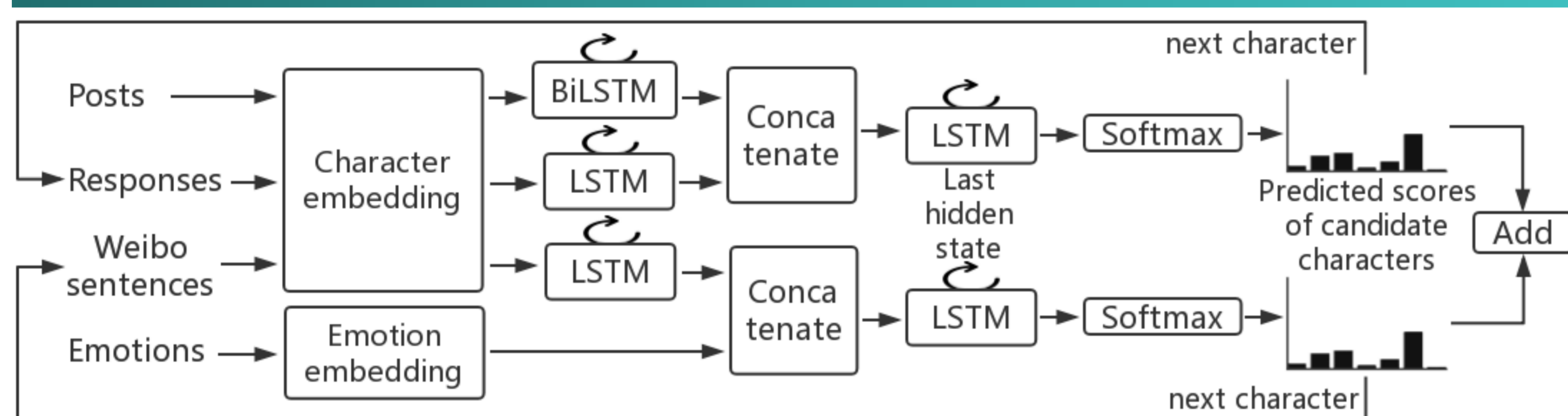


Fig.3 Structure and generation process of P2R&E2R model

## Preprocessing

1

- Removing post & response pairs without Chinese characters.
- e.g. -How are you?—Fine.

2

- Removing extra duplicate characters (3 times at most).
- e.g. -哈哈哈哈哈!!!-

3

- Removing low-frequent (frequency < 50) characters.
- e.g. -这是饕餮。

## Results

	like		sad		disgust		anger		happy		total	
	1	2	1	2	1	2	1	2	1	2	1	2
Label 0	121	109	84	92	82	92	85	85	71	76	443	454
Label 1	11	24	31	40	105	82	110	107	36	25	293	278
Label 2	68	67	85	68	13	26	5	8	93	99	264	268
Average	0.735	0.790	1.005	0.880	0.655	0.670	0.600	0.615	1.110	1.115	0.821	0.814

Table.2 The evaluation results of our run submissions.

Post	为什么电影院不卖周边呢[眼泪] Why don't movie theaters sell peripheral products [tears]	
Response	1	2
Like	我也喜欢。 I also like.	因为我喜欢看电影 Because I like to watch movie
Sad	[泪]我也要 [tears] I also want it	因为我没看过[泪流满面] Because I also have not seen it [tears]
Disgust	这是什么情况啊? What's going on?	因为你没有看电影啊 Because you didn't watch the movie
Anger	什么意思?!	什么电影啊? What movie?
Happy	[偷笑] [laughing]	哈哈,好多人知道,不知道有什么好玩的? Haha, many people know, don't know what fun?

Table.3 The responses comparison of P&E2R and P2R&E2R model.

## Conclusions

We use a character-based preprocessing method after analyzing the dataset.

We propose two different methods to generate responses with emotions.

The P&E2R model is a method of concatenating emotion labels with posts.

The P2R&E2R model is a method by training posts and emotion labels respectively.

Both methods utilize the beam search to improve the performance of responses.

Both submissions get average scores above 0.8, proving that our methods are effective to some extent.

In the future, we will pay more attention to the diversity of response generation.

