All is for Making a Machine Talk More Like People

Di Wu

1. Preface

This research plan is focused on driving the behavior of dialogue system to approach human performance. Next sections describe original ideas inspired by recent work of dialogue system. All of these could be left as future work containing conception completion and code implementation.

2. Dialogue System with Personal Linguistic Style using Adversarial Learning

To the best knowledge of the author, there has been few attention focused on creating a generation-based dialogue system possessing personal linguistic style. Personal linguistic style is hard to be explicitly defined and represented because it is a global feature of one's speaking style. Hence, it is a question how the framework based on seq2seq, which generates one word by turns, could correctly understand this kind of global feature and generate a sentence with strong personal style. The widely used dialogue dataset obtained from social networking sites, such as Weibo, will fail in this task. These public conversations contain so much noise to extract distinctive personal linguistic styles because of the mixture of various personal styles. So a dataset which specifically contains the utterances of a few people with distinctive personality should suit better.

Take the popular TV drama "The Simpsons" as an example, assuming that we have one dataset named "Homer Zone". In "Homer Zone", the posts are all responded by the leading role Homer. He has very funny personality far different from the ordinary. Here only one-turn conservation is considered for simplicity. Based on encoder-decoder, the task is to encourage the dialogue system to generate Homer-like responses. In other words, the system should distinguish the speaking style of Homer from the ordinary. Inspired by (Li et al., 2017), the adversarial learning may help achieve this goal. The framework is as follows:

Generator: The generator is similar to seq2seq model which is pertained in widely used large-scale dialogue dataset for good generalization. It generates a response $\hat{y}$ give a post x by other characters. More advanced encoder-decoder framework could be adopted.

Discriminator: The discriminator is a binary classifier which takes a response y as input and outputs a label indicating the probability (scores) of this response being Homer-like (denoted as Q_+) or not (denoted as Q_-). It is pretrained using the responses from "Homer Zone" dataset and the public dataset as positive and negative samples, respectively. The input response y is encoded into a vector representation using a LSTM (or GRU) encoder, then fed into a 2-class softmax function.

Policy Gradient Training: The main idea is to encourage the generator to generate Homer-like responses which can fool the discriminator. During the training step of generator, the generated response $\hat{y}$ is fed into discriminator and the output of the discriminator is treated as a reward r (Q_+) for optimizing generator parameters. Then the policy gradient method is combined with adversarial learning as can be seen in eq(1).

$$\nabla J(\theta) \approx [Q_+ - b(\hat{y})] \nabla \sum_t \log p(\hat{y}_t | x, \hat{y}_{1:t-1}) \tag{1}$$

where p denotes the probability of the generated responses, $b(\hat{y})$ denotes the baseline value. During the training step of discriminator, its parameters are updated by taking the generated responses $\hat{y}$ and responses y by Homer as negative and positive samples, respectively. More details and tricks such as REGS (Reward for Every Generation Step) can refer to (Li et al., 2017).

For number of training iterations **do**

  For i=1,D-steps **do**

    Sample Y from "Homer Zone" dataset;

    Sample $\hat{Y}$ from public dataset;

    Update D using Y as positive samples and $\hat{Y}$ as negative samples;

  End

  For i=1,G-steps **do**

    Sample Y from "Homer Zone" dataset;

    Sample $\hat{Y}$ generate by G$(\cdot|X)$;

    Compute Reward r for $\hat{Y}$ using D;

    Update G on $(X, \bar{Y})$ using reward r;

  End

End

Fig. 1. A brief review of the proposed adversarial reinforcement algorithm for training the generator G and discriminator D.

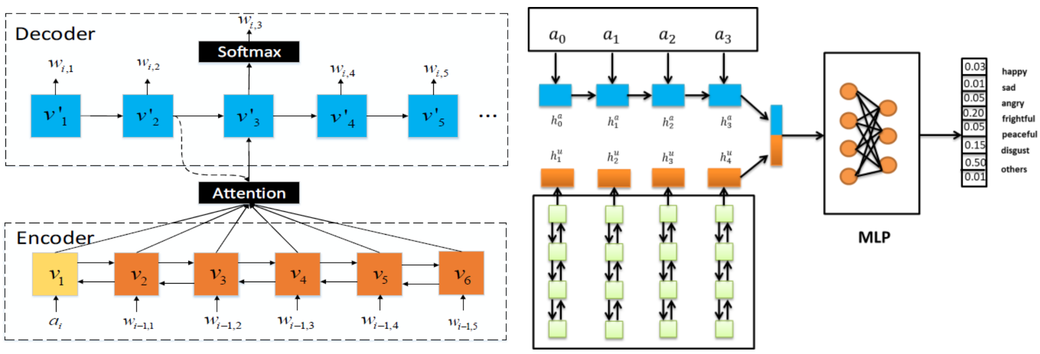
The method described above is a way which implicitly captures one's personal linguistic style by adversarial learning. It encourages the generator to generate responses $\hat{y}$ highly related to the target person (Homer). More explicit way such as directly embedding various linguistic personalities into dialogue system is still in thinking.

3. Dialogue System with Emotional Actions using Reinforcement Learning

(Zhou et al., 2018) proposed Emotional Chatting Machine (ECM) which meaningfully addressed the emotion factor in large-scale conversation generation. With embedding emotion categories, ECM introduces the two additional memory modules. The internal memory controls emotion expressing process during decoding, while the external memory controls the generation of emotion-related word. Because the main goal of ECM is generating an emotional response, it presents all kinds of emotional responses rather than specifically select one depending to the post. In fact, for the same post, different people have different emotions due to their different personalities. Hence, emotion-selecting task has practical significance. It can be seen as another aspect of learning a representation of ones' personality. Usually, people will not change their emotions frequently during conversation. To make sense, here the main goal is obtaining a chatbot with easily changed emotions. In other words, we attempt to model an interesting personality quite different from the ordinary.

Here "Homer Zone" dataset could be extended with emotional labels (actions). Because of drama characters' exaggerated personalities, their emotions change more frequently than the ordinary. Therefore, the dataset based on drama characters' stage lines will suit this task better. It is worth mentioning that (Xu et al., 2018) managed the flow of human-machine interactions with the dialogue actions as policies. By analogy, we can treat emotion categories defined in (Zhou et al., 2018) as emotion actions.

To the best knowledge of the author, the applications of reinforcement learning in NLP are usually improving the sentence generation quality such as informativity, coherence, and ease of answering as reported in (Li et al., 2016b). (Xu et al., 2018) also utilized reinforcement learning to make dialogue system verge to generate long and highly-relevant conversations. Towards these known contributions, reinforcement learning should also be able to help recognize deep patterns of people's behavior when applied in NLP. Here presents a framework equipped with reinforcement learning.



(a) Generate network (b) Policy network

Fig. 2. Policy network and generation network similar to (Xu et al., 2018)

"Homer-Zone" Dataset: This dataset $D = \{d_i\}_{i=1}^{N}$ where $d_i = \{(O_{i,1}, H_{i,1}, a_{i,1}), \ldots, (O_{i,n}, H_{i,n}, a_{i,n})\}$ refers to an Others-Homer dialogue with $O_{i,k}$ the kth post by others, $H_{i,k}$ the kth response by Homer and $a_{i,k}$ the labeled emotional action of $H_{i,k}$, respectively.

Generation Network: From fig. 2.(a), one could see that generation network is quite close to the standard encoder-decoder architecture with an attention mechanism (Bahdanau et al., 2015). The only difference is attaching emotional action a_i to the top of the long sentence as a special word. The architecture of encoder-decoder with internal memory and external memory (Zhou et al., 2017) could also be utilized as the generation network.

Policy Network: The architecture of policy network is shown in fig. 2.(b). The emotional actions $a_{0:t-1}$ of Homer are encoded into action embedding vectors by a GRU encoder, where a_0 is the initial state of emotional action. The input post by other characters at the corresponding next moment *t* is encoded by a bidirectional GRU encoder. Its last hidden state is concatenated with the action embedding vector described above. Then the concatenation $\{h_{t-1}^{a}; h_t^{w}\}$ is fed into a MLP with a softmax layer. The final output is a probability distribution of the emotional action a_t at the moment *t*.

Reinforcement Learning: The policy network is trained using reinforcement learning. The history buffer $H = \{h_i\}_{i=1}^{N}$ where $h_i = \{(O_{i,1}, a'_{i,1}), \ldots, (O_{i,n}, a'_{i,n})\}$ refers *i*th history track. $O_{i,k}$ denotes the *k*th post by other characters while $a'_{i,k}$ denotes the emotional action selected by the policy network. The reward is defined as the summation of scores valuating h_i:

$$r(h_i) = \sum_i^N \sum_j^n \mathrm{scores}(a'_{i,j}, a_{i,j}) \tag{2}$$

where $a_{i,j}$ is the true action taken by Homer.

Driven by the reinforcement learning, this policy network could be expected to recognize the emotion action pattern of Homer. The selected emotional action will fed into generation network as part of input. Then generation network will output Homer-likes response with Homer-like emotions.

4. Dialogue System with User Memory

It is challenging to endow a dialogue system with an identity. (Li et al., 2016) first addressed embedding persona to a conversation model. Then (Qian et al., 2017) proposed an elaborate framework with tricks such as profile detector, bidirectional decoder and position detector. These contributions are focused on embedding identity into a dialogue system itself. To possess further intelligence, a dialogue system should also remember the identity of its service target. This additional function could be able to notably boost user experience.

The task of memorizing the identity information of users is quite close to the task-oriented dialogue systems supposed to understand the intention of users. (Zhang et al., 2018) proposed a novel memory-augmented dialogue management model (MAD) which employs a memory controller with a slot-value memory and an external memory. Their work found a quite reasonable way to dynamically capture user information and update the internal memory of the dialogue system during conversation. Hence, dialogue system with user memory can be seen as an extension of their work in open-domain conversation generation by simply substituting the slot type. Combined with the method proposed by (Qian et al., 2017), reasonable response can be generated by extracting essential user information from internal memory. The whole conversation scenario can be seen as fig. 3.

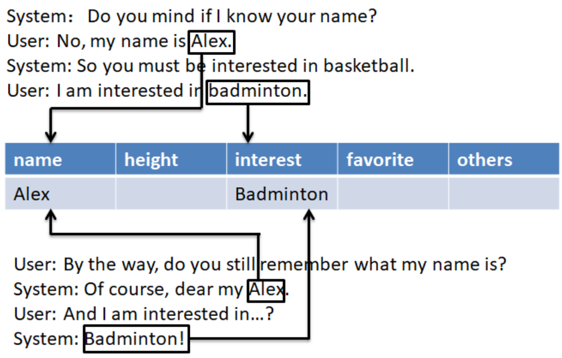


Fig. 3. Demonstration of the dialogue system with user memory

5. Combination of Computer Vision and Natural Language Processing via Emotions

Human beings have many ways to express their feelings. Besides directly speaking out what we think, what appear in our face also shares our feeling with others. For example, people are used to be regarded as in a happy mode when they smile without saying words like “I am feeling happy”. We could even speculate that a man will speak positive words when we see him in a positive mode. Hence, vision and language can be naturally connected via human emotions.

Back to the scenario of AI, the techniques of computer vision can easily capture the features of people’s emotional appearance, such as the curvature of the mouth or the shape of eyes. At the meantime, it is a bit more complex for natural language processing to recognize emotional patterns of contexts because of lacking of intuition. So are the emotional features extracted by computer vison able to help understand the emotional features of natural language? Assuming we have a large-scale dataset in which emotional appearance are labeled with corresponding emotional words. Will it work if we use this dataset to train a model which takes a photo as input and predicts words said by the character inside? I think it is valuable to research this problem which is quite different from the caption problem in computer vision. If one day it comes true, we could imagine that a NPC character will smile to the player, saying “I feel better, thank you for your help!” or “You should not do this, it is not good!" with angry eyes—all depend on his internal emotion system, i.e. his personality!

And all is for making a machine talk more like people.

6. References

Can Xu, Wei Wu, Yu Wu. 2018. Towards explainable and controllable open domain dialogue generation with dialogue acts. ICLR.

Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. 2015. Neural machine translation by jointly learning to align and translate. ICLR.

Jiwei Li, Will Monroe, Tianlin Shi, Alan Ritter, and Dan Jurafsky. 2017b. Adversarial learning for neural dialogue generation. EMNLP.

Jiwei Li, Will Monroe, Alan Ritter, Michel Galley, Jianfeng Gao, and Dan Jurafsky. 2016b. Deep reinforcement learning for dialogue generation. arXiv preprint arXiv:1606.01541.

Jiwei Li, Michel Galley, Chris Brockett, Georgios P. Spithourakis, Jianfeng Gao, Bill Dolan. 2016a. A Persona-Based Neural Conversation Model. ACL.

Hao Zhou, Minlie Huang, Tianyang Zhang, Xiaoyan Zhu, Bing Liu. 2018. Emotional chatting machine: emotional conversation generation with internal and external memory. AAAI.

Qiao Qian, Minlie Huang, Haizhou Zhao, Jingfang Xu, Xiaoyan Zhu. 2018. Assigning Personality/Identity to a Chatting Machine for Coherent Conversation Generation. IJCAI-ECAI.

Zheng Zhang, Minlie Huang, Zhongzhou Zhao, Feng Ji, Haiqing Chen, Xiaoyan Zhu. 2018. Memory-augmented dialogue management for task-oriented dialogue systems. arXiv preprint arXiv:1805.00150.