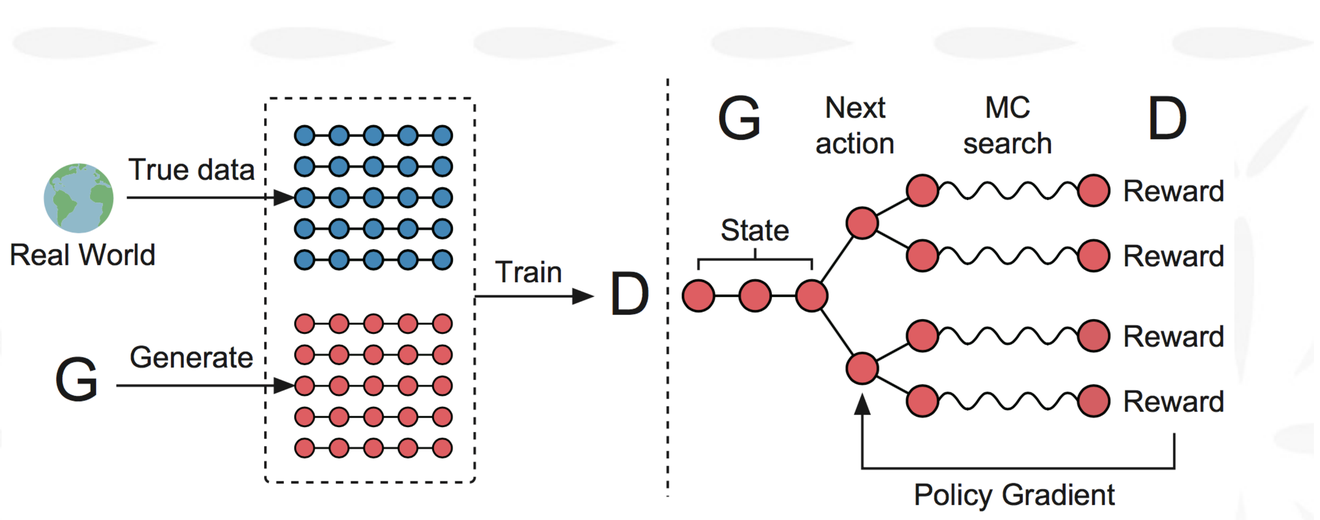
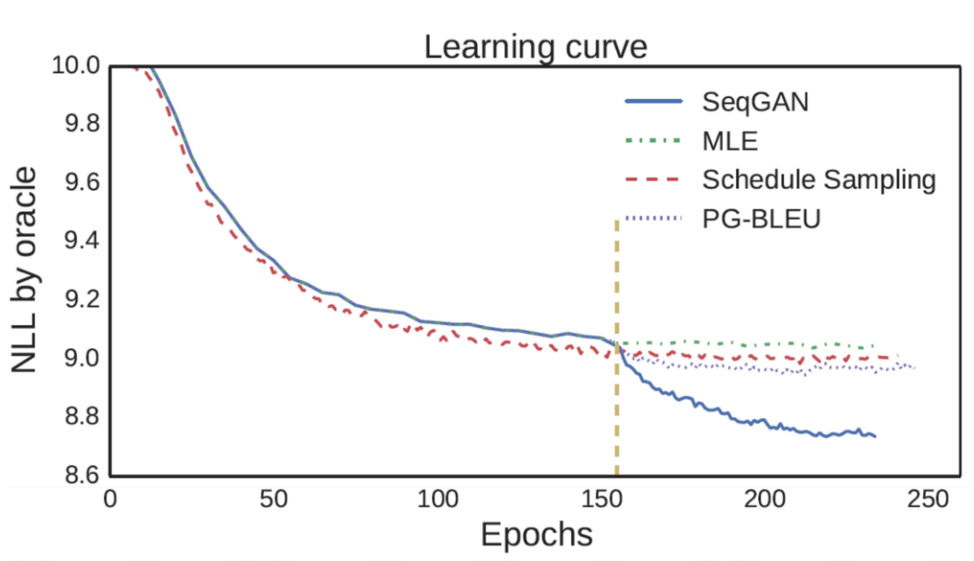
首届致远学术节
学生科研成果展示

Sequence Generative Adversarial Nets with Policy Gradient

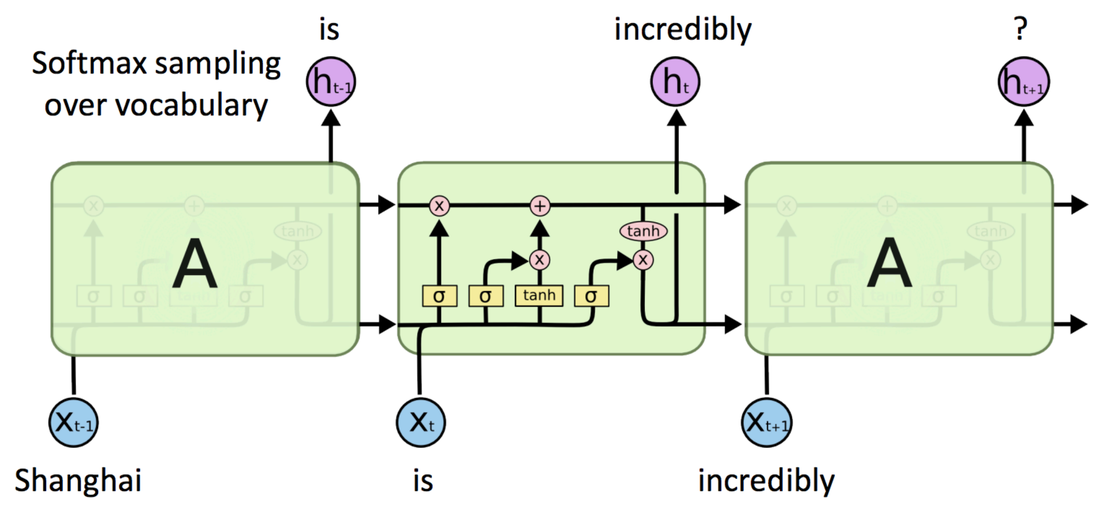
Lantao Yu, Weinan Zhang, Jun Wang, Yong Yu

The Thirty-First AAAI Conference on Artificial Intelligence

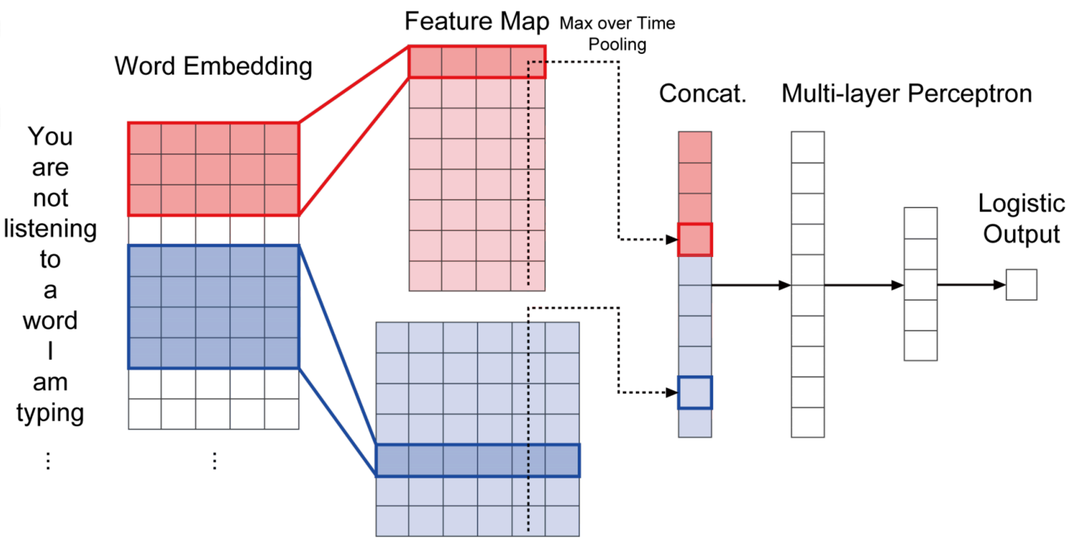




Generative Adversarial Network (GAN) has been a great success in generating natural looking images. However, the original GAN framework is limited to generate continuous data because the gradient computation involves backpropagation through generator's output. In this work, we proposed a novel framework SeqGAN, which elegantly combines adversarial training with policy gradient algorithm to generate discrete sequential data (*e.g.* natural language). So far many variants of SeqGAN have been proposed and successfully applied to a wide variety of tasks including dialogue systems (Li *et al.,* 2017), machine translation (Yang *et al.,* 2017), image caption (Dai *et al.,* 2017), music generation (Lee *et al.,* 2017) and recommender systems (Yoo *et al.,* 2017).



G: Long Short-Term Memory Network



D: Text Convolutional Neural Network

Algorithm 1 Sequence Generative Adversarial Nets

Require: generator policy G_θ; roll-out policy G_β; discriminator D_ϕ; a sequence dataset $\mathcal{S} = \{X_{1:T}\}$

Initialize G_θ, D_ϕ with random weights θ, ϕ.
Pre-train G_θ using MLE on $\mathcal{S}$
$\beta \leftarrow \theta$
Generate negative samples using G_θ for training D_ϕ
Pre-train D_ϕ via minimizing the cross entropy
repeat
  for g-steps **do**
    Generate a sequence $Y_{1:T} = (y_1, \ldots, y_T) \sim G_\theta$
    for t in $1 : T$ **do**
      Compute $Q(a = y_t; s = Y_{1:t-1})$ by Eq. (4)
    end for
    Update generator parameters via policy gradient Eq. (8)
  end for
  for d-steps **do**
    Use current G_θ to generate negative examples and combine with given positive examples $\mathcal{S}$
    Train discriminator D_ϕ for k epochs by Eq. (5)
  end for
  $\beta \leftarrow \theta$
until SeqGAN converges

个人信息：Lantao Yu (于澜涛)
邮　　箱：lantaoyu@hotmail.com