Non-Monotonic Sequential Text Generation

Joint work with:
Sean Welleck, Kyunghyun Cho and Hal Daumé III
Sequential Text Generation

Word Descrambling:  How are you?
Sequential Text Generation

Word Descrambling:

**Source:** you How ? are  

**Target:** How are you?

```
you  how  ?  are  <eos>
```

\[
f(\cdot | X) \quad f(\cdot | X, \text{how}) \quad f(\cdot | X, \text{how,are}) \quad f(\cdot | X, \text{how,are,you}) \quad f(\cdot | X, \text{how,are,you,?})
\]

\[
\begin{array}{cccccc}
\text{How} \ (w_1) & \text{are} \ (w_2) & \text{you} \ (w_3) & ? \ (w_4) & <\text{eos}> \ (w_5)
\end{array}
\]

\[
\begin{array}{ccccccc}
s_1 & \rightarrow & s_2 & \rightarrow & s_3 & \rightarrow & s_4 & \rightarrow & s_5
\end{array}
\]
Sequential Text Generation

Word Descrambling:

**Target:** How are you?

Assume: Sentence order - $w_1 \ w_2 \ w_3 \ w_4 \ w_5$

generation order - $s_1 \ s_2 \ s_3 \ s_4 \ s_5$  **Monotonic**

**Question:** Can we do sequential text generation using a non-monotonic generation order? (i.e. sentence order and generation order is different)


**Imitation Learning**

*(Structured Prediction)*

**Target:** How are you?

**Goal:** Train $\pi$ to mimic $\pi^*$ using a loss function

**States:** $s_1 \rightarrow s_2 \rightarrow s_3 \rightarrow s_4$ **Monotonic**

**Actions:** you good ? bad orange other green words test are hi things How ..... 

**Transition:** $P(s'|s, \cdot)$ **Fixed**

**Policy:** $\pi(\cdot | s)$

**Oracle policy:** $\pi^*(\cdot | s)$ **Optimal Sequence of**

**Loss:** $D_{KL}(\pi(\cdot, s) \mid \mid \pi^*(\cdot, s))$
Imitation Learning

(Change State Space)

Binary Tree State Space

\[ s_1 \]

\[ s_5 \rightarrow s_6 \rightarrow s_7 \]

\[ s_2 \rightarrow \ldots \rightarrow s_3 \rightarrow s_4 \]
Oracle Polices:

Binary Tree State Space

\[ \pi^*_{Uniform} = \begin{cases} 
1, & \text{if } a = \text{<end>} \text{ and } Y_t = <> \\
\frac{1}{n}, & \text{n is the number of unique words in } Y_t \\
0, & \text{otherwise}
\end{cases} \]
Oracle Polices:

States: Not Monotonic
Transition: Not Fixed
Oracle policy: Optimal actions

Binary Tree State Space

\[ \pi^*_{\text{Uniform}} = \begin{cases} 
1, & \text{if } a = \langle \text{end} \rangle \text{ and } Y_t = <> \\
\frac{1}{n}, & \text{if } n \text{ is the number of unique words in } Y_t \\
0, & \text{otherwise} 
\end{cases} \]
Oracle Polices:

\[ \pi^*_{\text{Uniform}} = \begin{cases} 
1, & \text{if } a = \langle \text{end} \rangle \text{ and } Y_t = <> \\
\frac{1}{n}, & \text{n is the number of unique words in } Y_t \\
0, & \text{otherwise}
\end{cases} \]
Oracle Polices:

\[ \pi^*_{\text{Uniform}} = \begin{cases} 
1, & \text{if } a = \text{<end>} \text{ and } Y_t = \langle \rangle \\
\frac{1}{n}, & \text{n is the number of unique words in } Y_t \\
0, & \text{otherwise} 
\end{cases} \]

\[ \pi^*_{\text{coaching}}(\cdot \mid a) \propto \pi^*_{\text{uniform}}(\cdot \mid a) \pi(\cdot \mid a) \]

\[ \pi^*_{\text{anneal}}(\cdot \mid a) = \beta \pi^*_{\text{Uniform}}(\cdot \mid s) + (1 - \beta) \pi^*_{\text{coaching}} \]
Unconditional Generation
(Language Model)
Conditional Generation

(Descrambling)

**π* Samples**

- hey there, i should be!
- not much fun. what are you doing?
- not. not sure if you.
- i love to always get my nails done.
- sure, i can see your eye underwater while riding a footwork.

left-right

- i just got off work.
- yes but believe any karma, it is.
- i bet you are. i read most of good tvs on that horror out. cool.
- sometimes, for only time i practice professional baseball.
- i am rich, but i am a policeman.

uniform

- i do, though. do you?
- i like iguanas. i have a snake. i wish i could win. you?
- i am a homebody.
- i care sometimes. i also snowboard.
- i am doing okay. just relaxing, and you?
## Conditional Generation
*(Neural Machine Translation)*

| Oracle          | BLEU (BP) | Validation | Test  |       |       |       |       |       |
|-----------------|-----------|------------|-------|-------|-------|-------|-------|-------|
|                 |           | Meteor     | YiSi  | Ribes | Meteor| YiSi  | Ribes |
| left-right      | 29.47 (0.97) | 29.66      | 52.03 | 82.55 | 26.23 (1.00) | 27.87 | 47.58 | 79.85 |
| uniform         | 14.97 (0.63) | 21.76      | 41.62 | 77.70 | 13.17 (0.64) | 19.87 | 36.48 | 75.36 |
| + (end)-tuning  | 18.79 (0.89) | 25.30      | 46.23 | 78.49 | 17.68 (0.96) | 24.53 | 42.46 | 74.12 |
| annealed        | 19.50 (0.71) | 26.57      | 48.00 | 81.48 | 16.94 (0.72) | 23.15 | 42.39 | 78.99 |
| + (end)-tuning  | 21.95 (0.90) | 26.74      | 49.01 | 81.77 | 19.19 (0.91) | 25.24 | 43.98 | 79.24 |