Non-Monotonic Sequential Text Generation

Sean Welleck, Kianté Brantley, Hal Daumé III, Kyunghyun Cho

NYU · University of Maryland

github.com/wellecks/nonmonotonic_text

Introduction

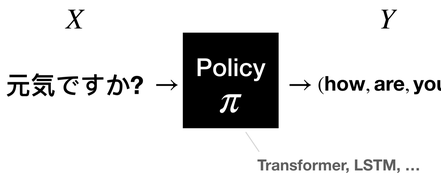
Can we train sequence models that *generate with learned, non-monotonic orders*?

In this work, we:

1. Propose a **Non-Monotonic Text Generation** framework for training text generation policies with **imitation learning**.
2. Define a class of oracles, including a **coaching oracle** that yields policies with **learned generation orders**.
3. Evaluate on **conditional** and **unconditional** generation tasks.

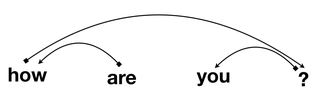
Sequential Text Generation

Sequentially generate tokens $Y = (y_1, \ldots, y_N)$.



Monotonic (e.g. left-to-right) generation

Non-monotonic generation



As Imitation Learning

- **Actions**: $a \in \mathcal{V} \cup \{\emptyset\}$ **word** or **end-token**.
- **State**: (x, T_t) where $\hat{T}_t = (a_1, ..., a_{<t})$ are predictions so far.
- **Goal**: Learn $\pi_\theta(a|s)$ that mimics an **oracle policy** $\pi^*(a|s)$.

Learning:

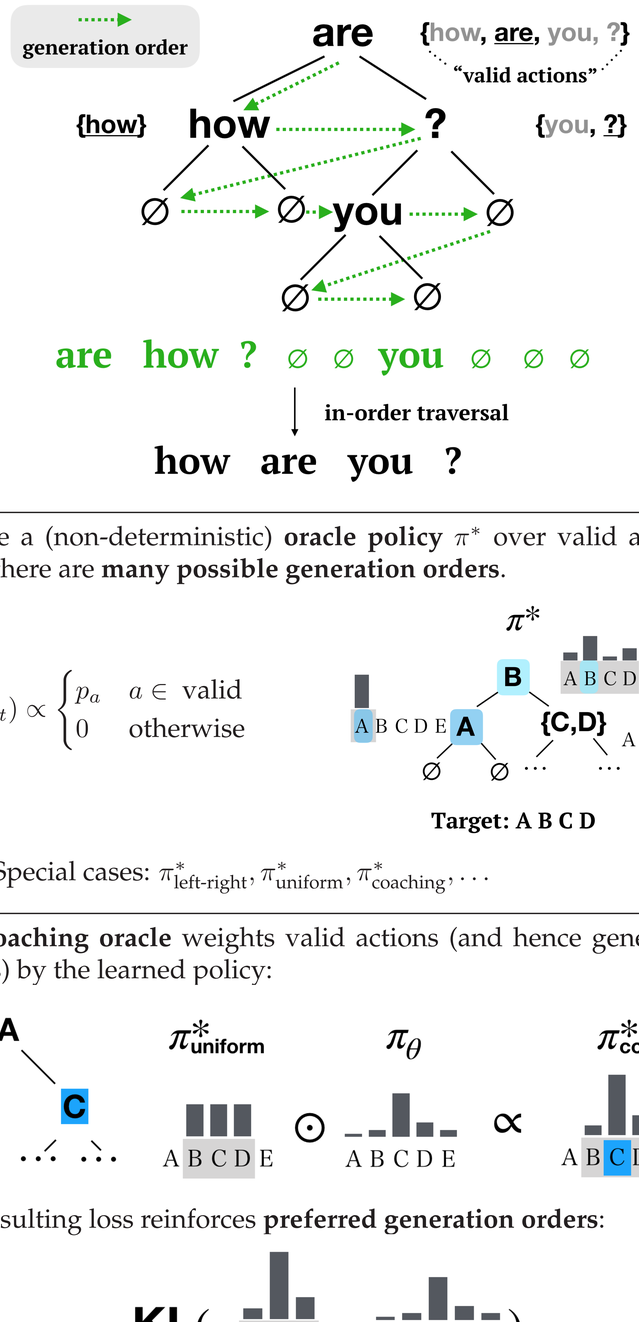
- Sample states, compute action costs with an **oracle policy** π^*
- Minimize action costs (**KL-divergence**) at sampled states

$$\mathbb{E}_{t\sim U[2|Y|+1]}\mathbb{E}_{s_t\sim d^t_{\pi^*}} D_{\mathrm{KL}}\left(\pi^*(\cdot|s_t)||\pi_\theta(\cdot|s_t)\right)$$

**See paper for a more general formulation.

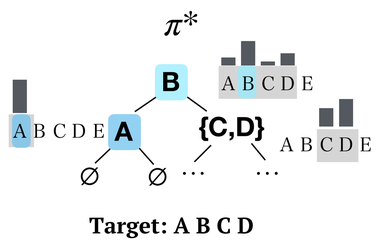
Non-Monotonic Text Generation

Generate a word at an arbitrary position, then recursively generate words to its left and then words to its right, yielding a **binary tree**.



Imitate a (non-deterministic) **oracle policy** π^* over valid actions, since there are **many possible generation orders**.

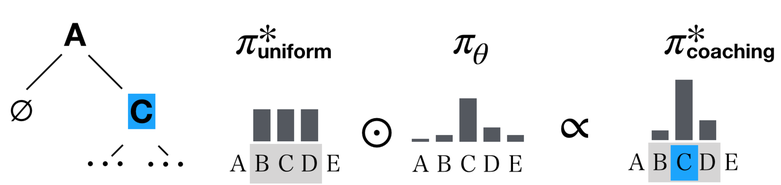
$$\pi^*(a|s_t) \propto \begin{cases} p_a & a \in \text{ valid} \\ 0 & \text{otherwise} \end{cases}$$



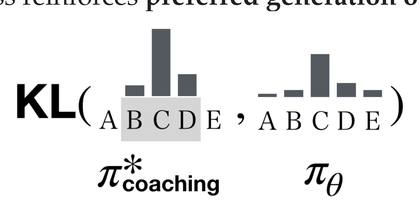
Target: A B C D

- Special cases: $\pi^*_{\text{left-right}}, \pi^*_{\text{uniform}}, \pi^*_{\text{coaching}}, \ldots$

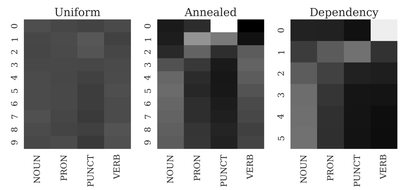
The **coaching oracle** weights valid actions (and hence generation orders) by the learned policy:



The resulting loss reinforces **preferred generation orders**:



Unconditional generation



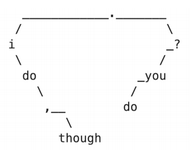
Oracle	%Unique	BLEU
left-right	17.8	47.0
uniform	98.3	40.0
annealed	93.1	56.2
validation	97.0	-

left-right

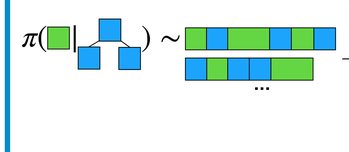
- hey there , i should be !
- not much fun . what are you doing ?
- not . not sure if you .
- i love to always get my nails done .
- sure , i can see your eye underwater while riding a footwork .

coaching (annealed)

- 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 ?



Variable-length Text Infilling



Initial Tree	Samples
	○ lasagna is my **favorite food !** ○ my **favorite food** is mac and cheese **!** ○ what is your **favorite food** ? pizza , i love it **!** ○ whats your **favorite food** ? mine is pizza **!** ○ seafood is my **favorite** . and mexican **food !** what is yours ?
	○ **hello** ! i like classical **music** . do you **?** ○ **hello** , do you enjoy playing **music ?** ○ **hello** just relaxing at home listening to fine **music** . you **?** ○ **hello** , do you like to listen to **music ?** ○ **hello** . what kind of **music** do you like **?**

Conditional Generation

Oracle	Validation BLEU	Validation F1	Validation EM	Test BLEU	Test F1	Test EM
left-right	46.6	0.910	0.230	46.3	0.903	0.208
uniform	44.7	0.968	0.209	44.3	0.960	0.197
annealed	46.8	0.960	0.230	46.0	0.950	0.212

Word Reordering (LSTM Policy)

Oracle	Validation BLEU (BP)	Validation Meteor	Validation YiSi	Validation Ribes	Test BLEU (BP)	Test Meteor	Test YiSi	Test Ribes
left-right	32.30 (0.95)	31.96	69.41	84.80	28.00 (1.00)	30.10	65.22	82.29
uniform	24.50 (0.84)	27.98	66.40	82.66	21.40 (0.86)	26.40	62.41	80.00
annealed	26.80 (0.88)	29.67	67.88	83.61	23.30 (0.91)	27.96	63.38	80.91
+tree-encoding	28.00 (0.86)	30.15	68.43	84.36	24.30 (0.91)	28.59	63.87	81.64
+⟨end⟩-tuning	29.10 (0.99)	31.00	68.81	83.51	24.60 (1.00)	29.30	64.18	80.53

Machine Translation (IWSLT De-En, Transformer Policy)