Learning from Easy to Complex:
Adaptive Multi-curricula Learning for Neural Dialogue Generation

[1]Hengyi Cai, [2]Hongshen Chen, [1]Cheng Zhang, [1]Yonghao Song, [1]Xiaofang Zhao, Yangxi Li, Dongsheng Duan, [2]Dawei Yin

[1]Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China; [2]JD.com

Motivation

Background:

- Current state-of-the-art neural dialogue systems are mainly data-driven and are trained on human-generated responses.
- Due to the subjectivity and open-ended nature of human conversations, the complexity of training dialogues varies greatly.
- The noise and uneven complexity of query-response pairs impede the learning efficiency and effects of the neural dialogue generation models.

Research Questions:

1. Conversation complexity embodies multiple aspects of attributes. **How to quantify the dialogue complexity?**
2. Babies learn to converse in an easy-to-complex manner and dynamically adjust their learning focus. **How to enable the dialogue model imitating such learning behaviors?**

Data Analysis: Curriculum Plausibility (Q1)

What defines the dialogue complexity?

Complexity quantification using conversational attributes: **Specificity**, **Repetitiveness**, **Query-relatedness**, **Continuity**, **Model Confidence.**

Low Specificity	*Do you have any pets?* *I do not. Do you?*	**High Query-relatedness**	*I win competitions.* *What kind of competitions do you win?*
High Specificity	*What kind of books do you read?* *Tom clancy n some james patterson.*	**Low Query-relatedness**	*I love art projects, including photography.* *Definitely. I am so tired though.*

Easy

Complex

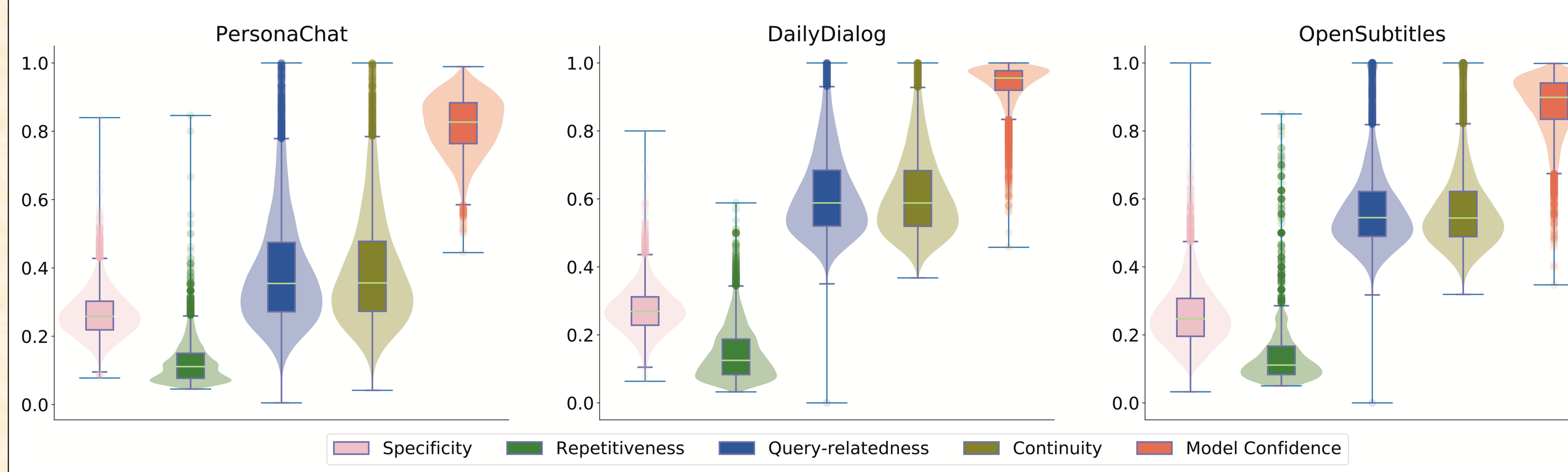
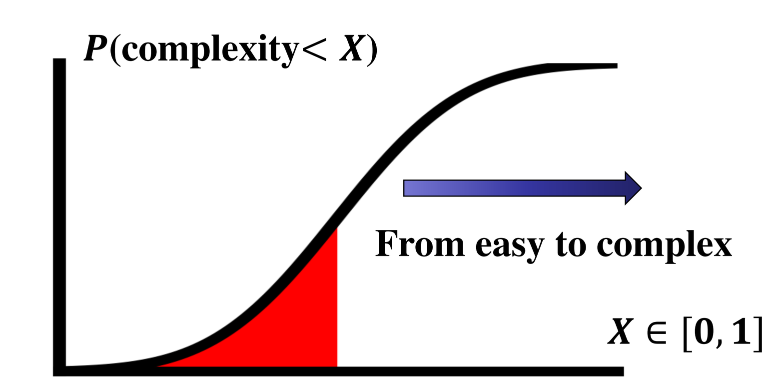


Figure 1: Violin plot with whiskers regarding five conversation attributes in three datasets.

- Dataset**:** PersonaChat [Zhang et al. 2018a], DailyDialog [Li et al. 2017], OpenSubtitles [Lison and Tiedemann 2016].
- Outliers frequently appear among all the distributions, which exhibits the uneven dialogue complexity.
- These attributes show little correlations with each other (Kendall correlations among these conversational attributes are near 0).

Single Curriculum Dialogue Learning (Q2)



Complexity is one of [Specificity, Repetitiveness, Query-relatedness, Continuity, Model Confidence].

- The curriculum is arranged by sorting dialogue training set according to the corresponding attribute.
- Progressing function: $f(t) \triangleq min(1, \sqrt{t\frac{1-c_0^2}{T} + c_0^2})$
- At training time step t, a batch of training examples is sampled from the top $f(t)$ portions of the total sorted training samples.
- T is the duration of curriculum learning and c_0 is set to 0.01.

- At the early stage, the model learns from samples drawing from the front part of the curriculum.
- As the advance of the curriculum, the difficulty gradually increases, as more complex training examples appear.

After training T batches, each batch of training instances is drawn from the whole training set, which is same as the conventional training procedure without a curriculum.

Adaptive Multi-curricula Learning (Q2)

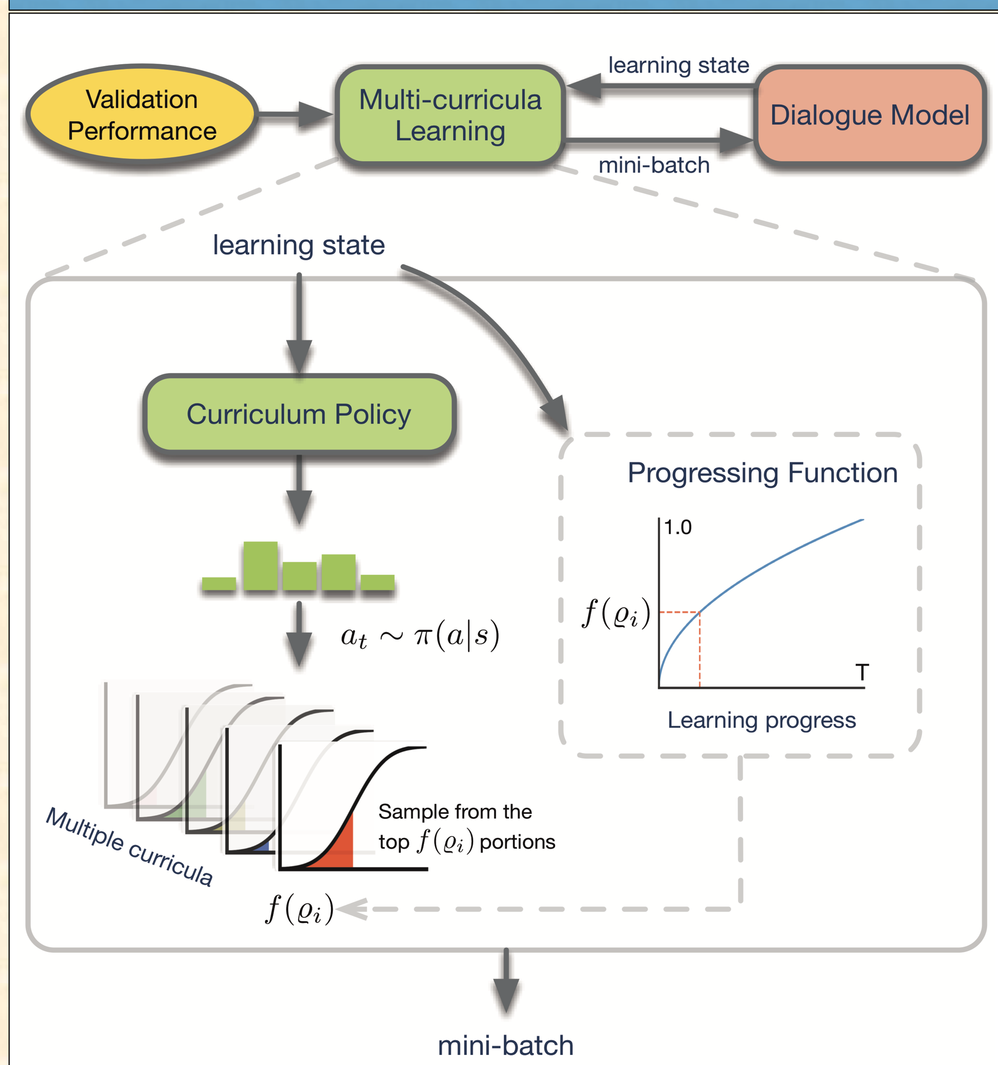


Figure 2: Overview of the proposed adaptive multi-curricula learning framework for neural dialogue generation. At training step t, the curriculum policy chooses one of the curricula to learn and the progressing function defines the learning progress on the selected curriculum.

- Dialogue complexity consists of multi-perspectives of attributes.
- Humans usually adjust their learning focus of multiple curricula dynamically in order to acquire a good mark.

We further introduce an adaptive multi-curricula learning framework, **to automatically choose different curricula** at different learning stages according to the learning status of the neural dialogue generation model.

- We provide the model with five different curricula, where each curriculum is prepared by ordering training set w.r.t. corresponding attribute metric accordingly.
- Scheduling mechanism acts as the **policy** π.
- **State**: the learning status of the dialogue model, including passed mini-batch number, the average historical training loss, etc.
- **Reward** R: The ratio of two consecutive performance deviations on validation set.
- **Action** $a_t \in \{0, 1, \cdots, k-1\}$ chooses one of the curricula, $k = 5$.
- Maximizing: $J(\theta) = \mathbb{E}_{\pi_\theta(a|s)}[R(s,a)]$.

Experiments

Metrics:

- **Dist-n** measures the ratio of unique n-grams to the total number of n-grams in a set of responses [Li et al., 2016];
- **Intra-n** measures the ratio of unique n-grams within each response [Gu et al. 2019];
- Embedding **Avg**, **Ext**, **Gre** metrics measuring the similarity between response and target word embeddings [Liu et al., 2016];
- **Coh**: Similarity between input and response word embeddings [Xu et al., 2018];
- **Ent-n**: n-gram entropy of responses [Serban et al., 2017].

Experimental Models:

(1) **SEQ2SEQ**: a sequence-to-sequence model with attention mechanisms (Bahdanau, Cho, and Bengio 2015), (2) **CVAE**: a conditional variational auto-encoder model with KL-annealing and a BOW loss (Zhao, Zhao, and Esk´enazi 2017), (3) **Transformer**: an encoder-decoder architecture relying solely on attention mechanisms (Vaswani et al. 2017), (4) **HRED**: a generalized sequence-to-sequence model with the hierarchical RNN encoder (Serban et al. 2016), (5) **DialogWAE**: a conditional Wasserstein auto-encoder, which models the distribution of data by training a GAN within the latent variable space (Gu et al. 2019).

	Models	BLEU	Dist-1	Dist-2	Dist-3	Intra-1	Intra-2	Intra-3	Avg	Ext	Gre	Coh	Ent-1	Ent-2
(a)	SEQ2SEQ	0.316	0.3967	2.190	5.026	77.24	87.00	90.73	58.85	47.22	65.91	62.87	6.674	10.368
	SEQ2SEQ (▲)	0.352	0.5406	3.537	8.343	82.74	92.28	95.47	62.20	47.57	67.07	66.87	6.875	10.723
	CVAE	0.290	0.5330	3.228	7.715	85.30	94.69	96.73	61.99	47.12	66.68	65.27	6.900	10.738
	CVAE (▲)	0.321	0.6530	4.572	11.326	89.39	96.90	98.28	63.08	47.37	67.09	66.81	6.973	10.866
	Transformer	0.195	0.7667	3.264	6.262	83.11	93.82	96.48	59.53	44.99	65.57	62.48	7.169	11.232
	Transformer (▲)	0.329	0.8468	4.291	8.829	89.39	97.92	99.29	62.33	46.24	66.54	65.35	7.183	11.331
	HRED	0.272	0.8109	4.217	8.948	84.25	95.20	97.19	60.63	45.91	66.33	63.53	7.082	11.149
	HRED (▲)	0.308	0.8926	5.332	12.281	91.45	97.89	98.93	62.25	46.53	66.53	65.22	7.156	11.274
	DialogWAE	0.124	0.9594	5.153	11.483	94.35	98.04	98.54	58.98	43.53	63.66	60.93	7.424	11.696
	DialogWAE (▲)	0.171	1.1388	6.890	15.842	96.65	99.41	99.68	63.81	45.90	65.63	65.63	7.462	11.845
(b)	SEQ2SEQ	0.399	1.542	9.701	22.005	91.10	96.97	98.15	67.50	47.41	68.45	68.39	6.933	10.921
	SEQ2SEQ (▲)	0.617	1.846	11.665	25.918	93.28	98.16	99.00	67.75	47.57	68.91	68.94	7.041	11.164
	CVAE	0.406	1.615	11.187	26.588	90.56	97.48	98.70	67.76	46.82	68.90	67.77	7.124	11.308
	CVAE (▲)	0.691	1.890	13.125	30.793	94.48	98.88	99.47	67.81	47.36	69.00	68.00	7.139	11.453
	Transformer	0.412	2.617	13.212	25.175	90.50	96.53	97.92	65.82	46.01	67.86	66.03	7.192	11.309
	Transformer (▲)	0.8063	2.917	15.509	30.954	94.38	98.59	99.26	66.52	46.79	68.40	66.65	7.307	11.651
	HRED	0.1746	2.323	11.563	22.471	94.01	98.45	99.30	65.09	45.91	67.49	65.09	7.141	11.331
	HRED (▲)	0.3834	2.448	12.880	26.355	94.18	98.65	99.36	65.37	46.43	68.14	65.22	7.058	11.341
	DialogWAE	0.0303	2.244	12.340	26.109	92.98	98.02	98.78	64.19	42.03	65.52	64.31	7.420	11.954
	DialogWAE (▲)	0.0814	2.654	16.311	36.591	92.79	98.73	99.53	65.27	43.41	66.60	65.62	7.539	12.106
(c)	SEQ2SEQ	0.140	0.3053	2.472	6.377	95.94	97.37	98.34	54.71	49.03	62.87	59.09	6.226	9.516
	SEQ2SEQ (▲)	0.172	0.4870	4.514	12.319	96.67	98.16	98.76	55.87	49.13	63.78	62.65	6.353	10.236
	CVAE	0.0522	0.3028	2.614	7.574	95.12	97.32	98.19	56.17	47.70	63.10	58.85	6.156	9.460
	CVAE (▲)	0.0429	0.4061	3.928	12.676	96.11	98.09	98.99	57.06	47.85	63.44	60.82	6.463	10.442
	Transformer	0.072	0.3883	1.737	3.503	95.38	97.13	98.18	55.10	48.16	62.69	57.45	6.661	10.362
	Transformer (▲)	0.050	0.5655	3.079	7.005	97.15	98.39	99.11	55.63	48.17	63.16	59.19	6.666	10.715
	HRED	0.0498	0.3311	1.900	4.465	95.34	97.38	98.15	55.41	48.34	62.79	58.92	6.346	9.715
	HRED (▲)	0.0795	0.6982	4.224	9.933	97.43	98.68	99.20	55.89	48.64	63.53	59.55	6.510	10.409
	DialogWAE	0.0038	0.4808	3.870	11.856	86.91	93.88	97.93	51.59	43.40	56.23	51.96	5.633	8.559
	DialogWAE (▲)	0.0352	0.7360	6.549	18.881	94.92	97.10	98.14	54.73	47.84	63.52	58.81	6.7859	11.187

Table 1: Automatic evaluation results (%) on three datasets: (a) PersonaChat, (b) DailyDialog and (c) OpenSubtitles. "▲" denotes training with our proposed framework.

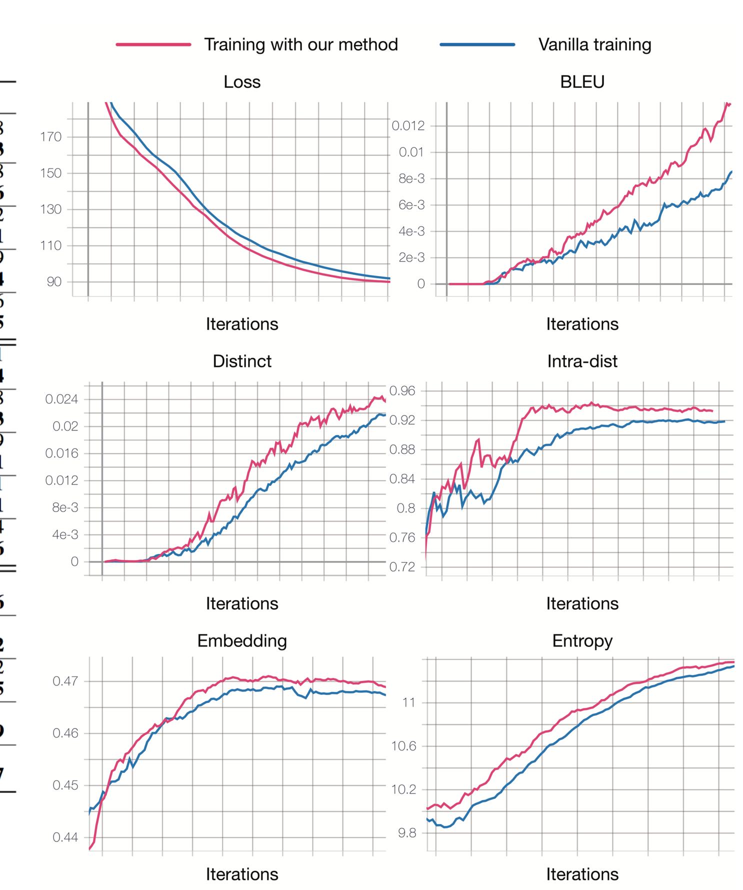


Figure 3: Vanilla mini-batch training *vs.* adaptive multi-curricula learning.

Examples with top learning frequencies	
Context:	Have you ever been to America?
Response:	No, but I'd really like to.
Context:	Can i ask you a question?
Response:	Sure, what do you want to know?
Context:	A new film is on. Have you ever seen it?
Response:	What kind of movie is it?
Examples with tail learning frequencies	
Context:	Ma'am?
Response:	May I have a magazine or something?
Context:	What took you so long?
Response:	Gross. Diarrhea?
Context:	I need more water!
Response:	oooo... gross... You're sweating all over the place.

Table 2: Examples with top and tail learning frequencies of SEQ2SEQ, using the proposed learning framework.

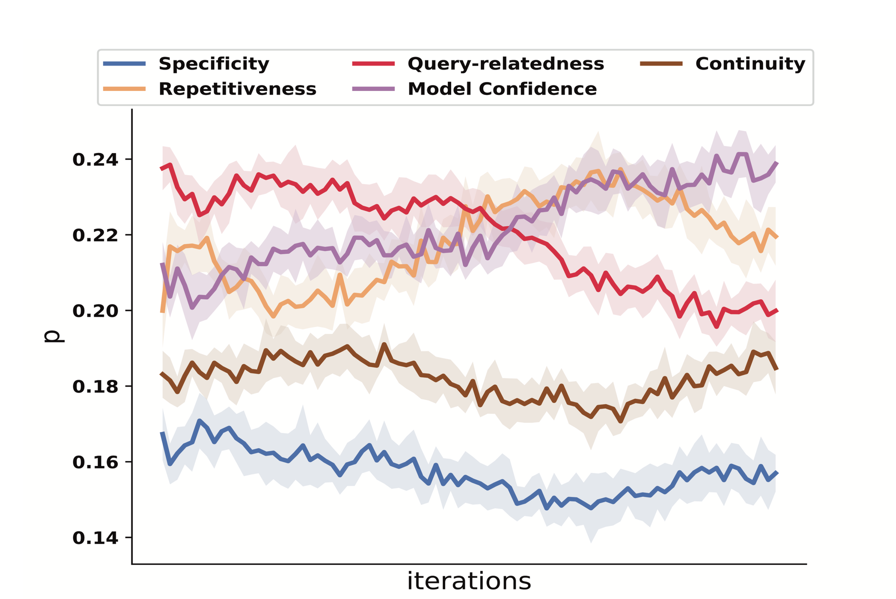


Figure 4: Trajectory of the curriculum learning action distribution $\pi(a_t|s_t)$ during the learning process. The model focuses on different curricula at different learning stages.