ECO Decoding: Entropy-Based Control for Controllability and Fluency in Controllable Dialogue Generation

Anonymous ACL submission

Abstract

Controllable Dialogue Generation (CDG) enables chatbots to generate responses with desired attributes, and weighted decoding methods have achieved significant success in the CDG task. However, using a fixed constant value to manage the bias of attribute probabilities makes it challenging to find an ideal control strength that satisfies both controllability and fluency. To address this issue, we propose **ECO decoding** (**E**ntropy-based **CO**ntrol), which dynamically adjusts the control strength at each generation step according to the model's entropy in both the language model and attribute classifier probability distributions. Experimental results on DailyDialog and MultiWOZ datasets show that our method achieves improved control accuracy while maintaining fluency and grammar, outperforming previous decoding methods across various models and settings. Furthermore, ECO decoding alleviates probability interpolation issues in multi-attribute generation, demonstrating its robust performance in both single and multi-attribute scenarios.

1 Introduction

Recently, Controllable Dialogue Generation (CDG) (Zhang et al., 2023; Zeng et al., 2023) has been proposed to enhance the realism and accuracy of responses generated by conversational models, improving the user experience. CDG enables chatbots to generate responses tailored to desired attributes like emotion and dialog-act.

In the field of controllable generation, training-based methods such as alignment tuning and **weighted decoding** approaches (Yang and Klein, 2021; Arora et al., 2022) have achieved notable success. While alignment tuning requires re-training large models from scratch, weighted decoding can be applied at inference time by multiplying the attribute probabilities from a trained classifier with

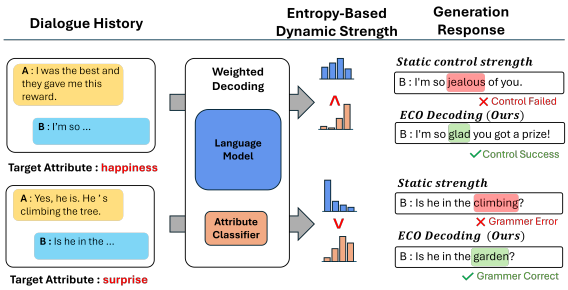

Figure 1: Controllable Dialog Generation method based on dynamic weighting with Eco Decoding. By dynamically determining the weights between the language model probability distribution and the attribute control probability distribution, it is possible to perform attribute control while maintaining fluency.

the language model distribution. This design enables controlled response generation with minimal additional data and training overhead, making weighted decoding an appealing strategy for CDG. Consequently, we focused on this weighted decoding strategy to effectively generate controllable responses.

In weight decoding methods, generating responses controlled by desired attributes involves the adjustment of the next token probability distribution modeled by the language model. This is achieved by multiplying the attribute probability of the generated response obtained from the attribute classifier and the next token probability. In this process, the control strength is used as the exponent of the attribute probability to control attribute bias. As the control strength increases, the generated tokens become more dependent on the token rank of attribute probability.

Multiplying the attribute probability alters the probability distribution of the language model, which can affect language modeling performance. When static control strength is used, the same control probability is continuously reflected in the generated sentence, even if the sentence has already

received sufficient attribute control or if specific words need to be generated for fluency. This can lead to a trade-off between controllability and fluency. Furthermore, the fact that the appropriate control strength varies depending on the situation is an important issue. If this is not properly accounted for, it can lead to decreased efficiency. Figure 1 shows an example of a failed response generation with these fixed static control strength.

To address this, we propose **ECO** (**E**ntropy-based **CO**ntrol strength) **Decoding**, which dynamically adjusts control strength at each decoding step. Specifically, we compute the entropy (Shannon, 2001) of both the language model's probability distribution and the attribute classifier's distribution at every step. If the language model assigns high probability (low entropy) to a particular token, we prioritize the model's prediction to preserve fluency. In contrast, when the model exhibits low confidence (high entropy), we increase the relative weight of the attribute probabilities to ensure stronger attribute control. Notably, our analysis result using *part-of-speech (POS)* in Appendix A.1 demonstrates that function words with fixed grammatical roles often yield lower entropy, whereas content words (e.g., nouns, verbs, adjectives) show higher entropy. This aligns with our intuition that words with higher uncertainty provide more leeway for attribute manipulation without significantly compromising fluency.

This dynamic control method effectively balances the language model's fluency with the attribute classifier's controllability, thereby achieving an optimal trade-off between naturalness and the desired attribute expression in the final generated sentences. To validate our intuition, we experiment with three existing controllable generation models using the DailyDialog (Li et al., 2017) dataset and MultiWOZ (Budzianowski and Casanueva, 2018) dataset. Experimental results demonstrate that ECO decoding achieves high controllability while maintaining text fluency across all models.

Our main contributions are as follows:

1. We raise the issue of static control strength in existing weighted decoding methods and propose a dynamic control strength approach to generate responses with high controllability as well as maintain fluency.

2. We show that ECO Decoding can be generalized to multi-attribute scenarios, alleviating probability interpolation issues that are commonly to single-attribute weighted decoding methods.

3. Experimental results show that the ECO decoding method outperforms the existing weighted decoding methods for all existing controllable generation models.

2 Related Work

2.1 Weighted Decoding

Controllable dialogue generation aims to generate a response, $R = \{r_1, r_2, ..., r_N\}$, with desired attributes, given dialogue history h and attribute c, using a pre-trained auto-regressive model (e.g. GPT2, (Radford et al., 2019), DialoGPT (Zhang et al., 2020)). Emotion and dialog-act can be attributes for controllable dialogue generation.

To condition on attribute c, the response generation given a dialogue can be formulated as follows:

$$P(R|h,c) = \prod_{i=1}^{N} P(r_i|r_{<i}, h, c) \quad (1)$$

Using Bayesian factorization, $P(r_i|r_{<i}, h, c)$ can be converted into the following equation.

$$P(r_i|r_{<i}, h, c) \propto P(r_i|r_{<i}, h)P(c|r_{\leq i}, h)^{\lambda} \quad (2)$$

where the first term $P(r_i|r_{<i}, h)$ represents the next token probability modeled by a language model, and the second term $P(c|r_{\leq i}, h)$ represents the attribute probability of the generated response obtained from the attribute classifier. In addition, control strength λ is added to the exponential term of the attribute probabilities to control attribute bias.

When dealing with multi-attribute control, Equation 3 can be extended by introducing the product of multiple attribute classifiers, assuming that the attributes are conditionally independent:

$$P(r_i|r_{<i}, h, C) \propto P(r_i|r_{<i}, h) \prod_{c_j \in C} P(c_j|r_{\leq i}, h)^{\lambda} \quad (3)$$

where C denotes the set of target attributes. The product of probabilities is typically implemented as the sum of logits.

2.2 Weighted Decoding Models

FUDGE Yang and Klein, 2021 trained a classification model for partial sequences through an

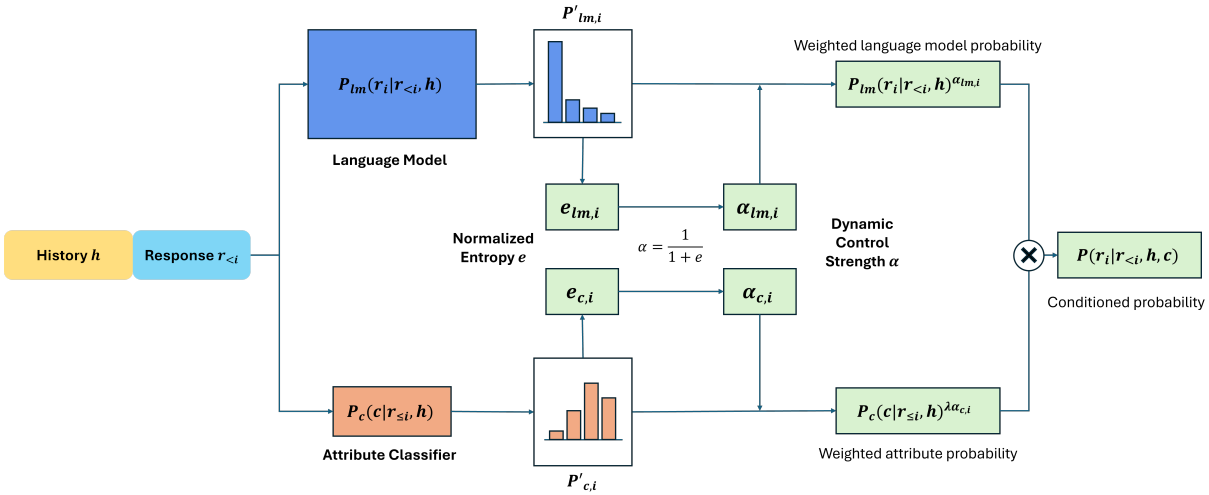

Figure 2: An illustration of controllable dialogue generation using the weighted decoding method, incorporating ECO Decoding.

external attribute classifier. Specifically, for each training example $\{(x, c)\}$, where x is sentence and c is class label, the classifier is trained on all partial sequences $\{(x_{1:i}, c)\}$ at each step. During inference, at a given time step i, the classifier predicts the probability that appending the top k candidate tokens to the generated text will satisfy the attribute c in future generations.

Director Arora et al., 2022 addressed the inefficiency issue of requiring a external model during inference. It integrates the language model and attribute classification functionality into a single model, overcoming the inefficiency of the external classifier evaluating the attribute for every candidate token. To address this issue, an additional classification head is introduced, which takes the last hidden state as input and computes the probability that each token in the vocabulary satisfies the specified attribute. This allows for the effective incorporation of attribute information without the need for a external classifier.

DASC Zhang et al., 2023 addressed the computational inefficiency issues arising from dual-head architectures. DASC introduces Attribute Token Embedding and Attribute Semantic Embedding concepts, employing a semantic space-based weighted decoding mechanism to reduce the number of parameters while improving computational efficiency. Each token is associated with an embedding that captures its attribute semantics, and these embeddings are projected into an attribute semantic space via attribute-specific linear layers. This design facilitates smooth control over multiple attributes and enables effective interpolation among attribute embeddings, allowing more diverse range of attribute combinations.

3 Methodology

3.1 Entropy-based Control Strength

The existing weighted decoding methods apply a fixed control strength and they are not flexible enough to handle situations where stronger or no more control is needed. In such cases, they may fail to control attribute, or even if they succeed, the fluency and grammar may degraded. To solve this problem, we propose the ECO decoding method that utilizes the entropy of the probability distribution to dynamically adjust the control strength. Dynamic control strength allows to achieve higher attribute control rates, while maintaining generation quality, including context and grammar.

Entropy is a measure of the uncertainty of a probability distribution, which is lower when the probability distribution is focused on a specific value and higher when it is more evenly distributed. Given this property, the higher the entropy of the next token probability distribution is, the more likely it is to contain a variety of plausible candidates. This is an advantageous property for exploring plausible options that satisfy desired attribute. Based on this insight, a novel mechanism of dynamically controlling strength is developed by weighting probability distributions from language models and it controls each property inversely to their entropy score. That is, distributions with lower uncertainty are more strongly reflected. Figure 2 shows how ECO decod-

ing is working by using dynamic control strength based on both of the language model entropy and the attribute entropy. ECO decoding can be applied to the existing weighted decoding methods and requires no additional modules or training.

Language Model Entropy Dynamic control strength $\alpha_{x,i}$ is separately calculated for i-th generation step, and it can have different values while a sentence is generated. To calculate control strength, we select the top-k candidate tokens. From the probability distribution $P_{lm,i}$ of the language model, we construct the set S, which consists of the k tokens with the highest probabilities. Let $P'_{lm,i}$ denote the partial probability distribution of top-k tokens in S.

$$P'_{lm,i} = \{P_{lm}(t|r_{<i}, h)|t \in S\} \tag{4}$$

To convert the partial probability distribution $P'_{lm,i}$ into a probability distribution, we recompute the probability distribution of the top-k tokens using a softmax function with temperature τ_{lm}.

$$e_{lm,i} = Entropy(Softmax(P'_{lm,i}/\tau_{lm})) \tag{5}$$

Attribute Entropy Weighted decoding methodologies for CDG utilize attribute classifier P_c to reflect attributes. For each candidate token t in the top-k token set S, concatenates the current sequence $r_{<i}$ with t and computes the probability $P_{c,i}([r_{<i}; t], h)$ which represents the probability of token t being part of the generated response while aligning with the target attribute to be controlled. The set $P'_{c,i}$ is the probabilities of the target attribute for all candidate tokens in top-k token set S. The attribute entropy $e_{c,i}$ is computed based on a probability distribution normalized by softmax the set of attribute probabilities $P'_{c,i}$ over τ_c, where τ_c is the attribute temperature for softmax.

Entropy Based Control Strength To assign higher weights to probability distributions with higher entropy, we utilize a control strength formula with an inverse function structure, as shown in Equation 6. The control strength $\alpha_{x,i}$ is applied to both the language model probability distribution P_{lm} and the attribute probability distribution P_c. The language model probability distribution and the attribute probability distribution are reflected by a power of their respective weight $\alpha_{x,i}$. The attribute probability distribution additionally reflects the strength scale factor λ. The value of λ allows to adjust whether to focus more on attribute control or language modeling performance. The final probability distribution for generating the next token $P(r_i|r_{<i}, h, c)$ is computed by multiplying the two weighted probability distributions as shown in Equation 7. If each of the control strength $alpha$ values were fixed at 1, the same result would be obtained as with the traditional weighted decoding methodologies.

$$\alpha_{x,i} = 1 + (\frac{1}{1 + e_{x,i}}) \tag{6}$$

$$\begin{aligned} P(r_i|r_{<i}, h, c) \propto P_{lm}(r_i|r_{<i}, h)^{\alpha_{lm,i}} \\ \times P_c(c|r_{\leq i}, h)^{\lambda * \alpha_{c,i}} \end{aligned} \tag{7}$$

3.2 Multiple Attribute Control Strength

Existing weighted decoding methodologies struggle to control multiple attributes simultaneously due to their fixed control strength. When using a fixed control strength for each attribute, the search space of attribute control strengths grows exponentially. Furthermore, even when control strength is applied, effectively incorporating more than two attributes remains a main challenge. In contrast, our proposed ECO-decoding method enables CDG to control generation by reformulating the final probability distribution based on multiple attributes. Dynamic control strength $\alpha_{x,i}$ adjusts the weight of probability distributions at each generation step based on the entropy of the language model and the entropy of each attribute, allowing more flexible and adaptive multi-attribute control. When C is the set of controlling attributes, the multiple attribute control formula for ECO-decoding is as follows:

$$\begin{aligned} P(r_i|r_{<i}, h, C) \propto P_{lm}(r_i|r_{<i}, h)^{\alpha_{lm,i}} \\ \times \prod_{c_j \in C} P_{c_j}(c_j|r_{\leq i}, h)^{\lambda * \alpha_{c_j,i}} \end{aligned} \tag{8}$$

$$P'_{c,i} = \{P_c([r_{<i}; t], h)|t \in S\} \tag{9}$$

$$e_{c,i} = Entropy(Softmax(P'_{c,i}/\tau_c)) \tag{10}$$

4 Experiments

4.1 Datasets

DailyDialog (Li et al., 2017) is an English open-domain dialogue dataset containing two main attributes (emotion, dialog-act). We treat each utterance as a response and its preceding utterances

Model	Accuracy	Rouge-1	Rouge-L	Dist-1	Dist-2	Grammar
			Emotion			
DialoGPT(176M)	-	9.00	8.53	0.58	0.76	90.21
FUDGE	76.98	9.06	8.60	0.60	0.75	90.30
+ ECO decoding	81.03 (+4.05)	9.13 (+0.07)	8.64 (+0.04)	0.62 (+0.02)	0.75 (-)	90.34 (+0.04)
Director	79.94	8.83	8.37	0.59	0.70	90.23
+ ECO decoding	82.82 (+2.88)	8.82 (-0.01)	8.34 (-0.03)	0.59 (-)	0.71 (+0.01)	90.30 (+0.07)
DASC	74.65	8.25	7.87	0.58	0.70	90.30
+ ECO decoding	75.74 (+1.09)	8.22 (-0.03)	7.79 (-0.08)	0.58 (-)	0.71 (+0.01)	90.39 (+0.09)
			Dialog-act			
DialoGPT(176M)	-	9.14	8.66	0.57	0.78	91.24
FUDGE	41.07	9.21	8.75	0.59	0.78	90.98
+ ECO decoding	46.42 (+5.35)	9.21 (-)	8.79 (+0.04)	0.62 (+0.03)	0.79 (+0.01)	91.00 (+0.02)
Director	70.96	10.43	9.94	0.62	0.78	91.18
+ ECO decoding	71.56 (+0.60)	10.46 (+0.03)	9.96 (+0.02)	0.63 (+0.01)	0.79 (+0.01)	91.15 (-0.03)
DASC	42.59	9.53	9.03	0.59	0.75	91.13
+ ECO decoding	47.17 (+4.58)	9.52 (-0.01)	9.05 (+0.02)	0.60 (+0.01)	0.76 (+0.01)	91.13 (-)

Table 1: Evaluation results for a single attribute of emotion or dialog-act on the DailyDialog test set. The scores in brackets indicate the performance gap between static control and dynamic control settings.

Model	Accuracy	Rouge-1	Rouge-L	Dist-1	Dist-2	Grammar
DialoGPT(176M)	-	6.72	6.19	0.53	0.44	86.42
			Emotion			
Director	73.99	6.42	5.90	0.57	0.44	86.36
+ ECO decoding	75.31 (+1.32)	6.54 (+0.12)	6.04 (+0.14)	0.55 (-0.02)	0.44 (-)	86.38 (+0.02)
			Dialog-act			
Director	67.30	5.84	5.41	0.62	0.57	86.47
+ ECO decoding	67.92 (+0.62)	5.88 (+0.04)	5.45 (+0.04)	0.63 (+0.01)	0.59 (+0.02)	86.43 (-0.04)

Table 2: Evaluation results for a single attribute of emotion or dialog-act on the MultiWOZ dataset.

as the dialogue history. The dataset provides four classes for dialog-act (inform, question, directive, commissive) and six classes for emotion (anger, disgust, fear, happiness, sadness, surprise), excluding "no emotion." It comprises 13,681, 882, and 1,286 examples for training, validation, and test, respectively.

MultiWOZ (Budzianowski and Casanueva, 2018) is a large-scale multi-domain dialogue dataset constructed from real human-to-human conversations, encompassing multi-turn dialogues across seven domains (restaurants, trains, attractions, hotels, taxis, hospitals, and police). Since the dataset itself does not provide labeled attributes, we employ an attribute classifier—used for evaluation—to label each utterance, and use these labels as the ground truth.

4.2 Experimental Settings

Language Model We employ DialoGPT (Zhang et al., 2020), pre-trained on a large-scale dialogue corpus, as our baseline. Most experiments use DialoGPT-small (176M), and we also evaluate DialoGPT-large (1.1B) for scalability. Additionally, we use Llama2-7B (Touvron et al., 2023) to test the general applicability of ECO decoding to large language models.

Weighted Decoding Methods We evaluate and compare the performances of ECO decoding with those of various controllable generation models with weighted decoding methods, including FUDGE (Yang and Klein, 2021), Director (Arora et al., 2022), and DASC (Zhang et al., 2023).

Implementation Details For the three weighted decoding method, the language model is frozen and each attribute classifier is trained on the training dataset. FUDGE is trained for 30 epochs with a batch size of 8 and a learning rate of 2e-5 for each attribute. For the Director, each attribute is fine-tuned for 20 epochs with a batch size of 32 and

Model	Accuracy(Emo)	Accuracy(Act)	Rouge-1	Rouge-L	Dist-1	Dist-2	Grammar
DialoGPT(176M)	-	-	9.00	8.53	0.58	0.76	90.21
FUDGE	66.17	44.17	8.21	7.82	0.57	0.74	90.20
+ ECO decoding	66.41 (+0.24)	45.57 (+1.40)	8.20 (-0.01)	7.81 (-0.01)	0.58 (+0.01)	0.74 (-)	90.21 (+0.01)
Director	80.48	60.65	9.41	8.99	0.58	0.73	90.22
+ ECO decoding	81.18 (+0.7)	61.20 (+0.65)	9.49 (+0.08)	8.97 (-0.02)	0.58 (-)	0.74 (+0.01)	90.23 (+0.01)
DASC	75.19	51.17	8.22	7.67	0.60	0.77	90.05
+ ECO decoding	77.22 (+2.03)	54.12 (+2.95)	7.60 (-0.62)	7.15 (-0.52)	0.61 (+0.01)	0.78 (+0.01)	90.19 (+0.14)

Table 3: Evaluation results for multiple attributes setting on the DailyDialog test set.

Model	**Accuracy(Emo)**	**Accuracy(Act)**	**Rouge-1**	**Rouge-L**	**Dist-1**	**Dist-2**	**Grammar**
DialoGPT(176M)	-	-	6.72	6.19	0.53	0.44	86.42
Director	63.62	49.29	7.22	6.72	0.44	0.48	86.38
+ ECO decoding	66.52 (+2.90)	52.59 (+3.30)	7.35 (+0.13)	6.72 (-)	0.45 (+0.01)	0.49 (+0.01)	86.41 (+0.03)

Table 4: Evaluation results for multiple attributes setting on the MultiWOZ dataset.

Model	Accuracy	Interest	Sensible
Emotion			
Director	2.82	2.96	2.75
+ ECO decoding	3.19 (+0.37)	3.16 (+0.20)	3.15 (+0.40)
Dialog-act			
Director	3.04	2.93	2.78
+ ECO decoding	3.42 (+0.38)	3.41 (+0.48)	3.36 (+0.58)

Table 5: Human Evaluation on DailyDailog test set (single attribute)

a learning rate of 1e-5. For DASC, each attribute is fine-tuned for 30 epochs with a batch size of 4 and a learning rate of 1e-5. All methods use greedy search (Li et al., 2016b), and the maximum sequence length is set to 128. All experiments are run on a single NVIDIA GeForce RTX 3090.

Our code is available at `https://anonymous.4open.science/r/ECO-46EF/`

4.3 Evaluation Metrics

Automatic Evaluation To assess *controllability*, we train two RoBERTa-based evaluators (Liu et al., 2019) on the DailyDialog training set: one for emotion and one for dialog-act. These evaluators achieve 89.66% and 80.60% accuracy on their test sets respectively and they are used to classify generated responses. Note that these evaluators are independent from the attribute classifiers in each weighted decoding method. For *quality*, ROUGE-1 and ROUGE-L (Lin, 2004) are measured by comparing generated responses to reference answers and Dist-1 and Dist-2 (Li et al., 2016a) are computed to evaluate diversity in the generated text. For *grammar*, the probability of grammaticality is utilized by the RoBERTa-based CoLA grammaticality model (Liu et al., 2019; Warstadt et al., 2019; Morris et al., 2020).

Human Evaluation Experiments on the Director model, which showed the best performance in emotion and dialog-act attributes, conducted human evaluation based on sampling 10 contexts for each attribute value from the test set. We evaluate our generated responses based on three aspects: (1) Accuracy: Response is generated according to the desired attribute. (2) Interest: Response is specific and creative enough to capture the user's attention, and avoids repetitive or generic outputs that may reduce engagement (e.g., repeatedly generating "That's great!" for the happy attribute). (3) Sensible: Response is grammatically correct and contextually coherent. We asked three expert evaluators to rate each metric on a scale of 1 to 5, with higher scores being better.

4.4 Experimental Results

Single Attribute Control Table 1 summarizes how effectively ECO decoding enhances controllability while maintaining fluency. We first run each model without attribute control to establish a baseline grammar score and then tune the control strength λ in each method (FUDGE, Director, DASC) until we match that baseline grammar score. Eventually, we evaluate the resulting Accuracy, Dist, and ROUGE metrics.

The results show that ECO decoding consistently improves the accuracy of both emotion and dialog-act attributes, surpassing static control methods. Notably, these improvements come with no drop in grammar quality, in contrast to many existing approaches. Furthermore, ECO decoding retains the Dist and ROUGE scores or it even slightly im-

Model	Accuracy	Rouge-1	Rouge-L	Dist-1	Dist-2	Grammar
DialoGPT(176M)	-	9.00	8.53	0.58	0.76	90.21
Director	79.94	8.83	8.37	0.59	0.70	90.23
+ ECO decoding	82.82 (+2.88)	8.82 (-0.01)	8.34 (-0.03)	0.59 (-)	0.71 (+0.01)	90.30 (+0.07)
DialoGPT(1.1B)	-	11.54	10.89	0.75	0.73	87.28
Director	75.66	11.76	11.15	0.74	0.73	87.18
+ ECO decoding	76.05 (+0.39)	11.82 (+0.06)	11.23 (+0.08)	0.74 (-)	0.73 (-)	87.25 (+0.07)
Llama2(7B)	-	15.23	12.99	0.35	0.08	90.60
Director	75.43	15.95	13.74	0.35	0.80	90.51
+ ECO decoding	75.66 (+0.23)	15.88 (-0.07)	13.63 (-0.03)	0.35(-)	0.80 (-)	90.55 (+0.04)

Table 6: Evaluation results for attributes of emotion on the DailyDialog test set with various size of model.

proves in some cases; we think that it confirms ECO decoding's ability to dynamically incorporate desired attributes without compromising fluency or overall response quality.

To assess generalizability, we further apply ECO decoding to the MultiWOZ dataset. As shown in Table 2, our method achieves similarly high controllability in this dataset while preserving grammar quality comparable to one of the DailyDialog dataset. These findings demonstrate the robustness of ECO decoding across different dialogue domains.

Multi Attribute Control Multi attribute control typically involves combining attribute probabilities via multiplication. Consequently, when interpolating across multiple distributions, differences in scale and calibration can make it difficult to maintain a proper balance, often leading to a decline in overall controllability compared to single-attribute control.

In Table 3, multi attribute control for the Emotion attribute achieves grammar performance on par with single-attribute control, yet exhibits a decrease in overall controllability. Conversely, multi-attribute control for the Dialo-act attribute appears to yield higher controllability relative to single attribute control. However, this does not necessarily indicate an actual improvement in controllability; rather, it likely reflects the selection of a relatively lower grammar score baseline due to differences in the experimental data. A same pattern is also observed in the MultiWOZ dataset experiments presented in Table 4.

By applying ECO decoding, we mitigate these interpolation issues and significantly enhance the controllability of both emotion and dialog-act attributes. Moreover, as in the single-attribute setting, dynamic weighting consistently maintains and even strengthens grammatical fluency and response quality in multi-attribute generation.

Language Model Scaling Table 6 compares the performance of the Director method and ECO decoding across models of varying sizes. From the smallest 176M parameter model to the 7B model, ECO decoding consistently achieves higher grammar scores while maintaining strong attribute controllability. This indicates that ECO decoding can be applied effectively to traditional weighted decoding methods regardless of model scale.

Human Evaluation Table 5 presents the human evaluation results. Consistent with the automatic metrics, ECO decoding substantially outperforms the baseline in generating coherent and attribute-aligned responses. Across both emotion and dialog-act attributes, Accuracy improves by about +0.38, Interest by +0.34, and Sensible by +0.49. These results suggest that ECO decoding not only better aligns generated responses with the desired attributes but also makes them more engaging and contextually coherent from a human perspective.

Robustness Test The control strength coefficient λ determines the proportion of weights in the probability distribution of the attribute classifier. Therefore, a larger lambda will tend to generate tokens that are more attribute-specific, resulting in a trade-off between increased attribute accuracy and decreased grammar score.

Figure 3 shows the results of applying traditional weighted decoding methodology and ECO-decoding for varying the control strength coefficient λ in a single attribute control setting. The experiments were conducted based on the Director and DASC method for two attributes, emotion and dialog-act, and the red line is for the Director and DASC. The green line is for the applica-

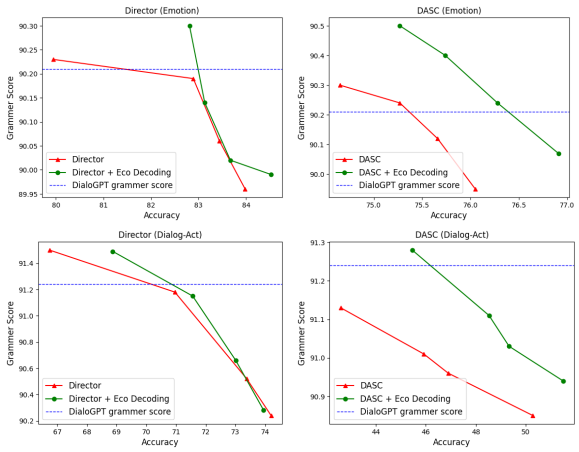

Figure 3: The single attribute control performance of the existing weighted decoding method (red) and ECO decoding (green) with respect to changes in the control strength λ. The y-axis represents grammar, and the x-axis represents accuracy. The blue dot line represents uncontrolled DialoGPT's grammar score.

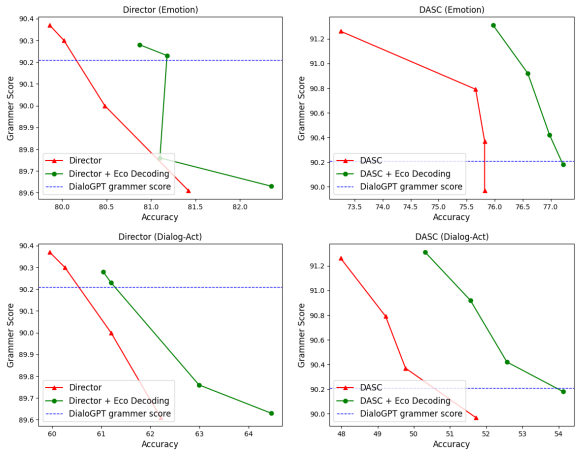

Figure 4: The multi attribute control performance of the existing weighted decoding method (red) and ECO decoding (green) with respect to changes in the control strength λ. The y-axis represents grammar, and the x-axis represents accuracy. The blue dot line represents uncontrolled DialoGPT's grammar score.

tion of ECO decoding with each method and the blue bashed line is the grammar score of vanila DialoGPT without attribute control according to each dataset. In all experiments, we observed a trade-off between grammar and accuracy, and showed that for the same grammar score, ECO decoding achieves higher attribute accuracy by dynamically applying weights. In other words, for the same control degree, ECO decoding produces higher quality responses. This demonstrates that our approach has a strong capability in controllable generation to maintain fluency while enhancing controllability, regardless of the λ values.

Figure 4 shows the performance by lambda in the multi attribute control setting. The results are the same for multi attribute as for single, with ECO decoding for each methodology resulting in higher controllability and grammar scores. The interesting thing is that the Director model is not a structured model for multi attribute control, and because of this, there is some performance variation by lambda.

In almost all cases using the DialoGPT-small model, we observed a trend of grammar score increase and then decrease as the lambda value increases. Due to the increase in grammaticality at low λ ranges, some experimental results showed higher grammar score with attribute control than DialoGPT without attribute control. This was not observed when using larger models such as Llama2, which could be thought of as a slight inconsistency in the performance of small language models. For fairness, all experimental results measured grammar score and attribute accuracy after the lambda value where the trade-off occurs.

5 Conclusions

We present ECO decoding as an entropy-based approach for dynamically adjusting control strength in weighted decoding. Unlike prior methods relying on static coefficients, ECO decoding utilizes the uncertainty of each attribute classifier during inference to improve controllability without sacrificing fluency.

ECO decoding is easily applicable to existing controllable generation frameworks because it does not need to train a generator directly. It effectively generalizes across single- and multi-attribute settings, and properly addresses interpolation issues in multi-attribute control by leveraging attribute-specific entropy.

Both automatic and human evaluations demonstrate that ECO decoding consistently outperforms static baselines across a range of models and control strengths, including strong performance on large-scale models like Llama2-7B. These findings highlight that ECO decoding is a effective solution for more precise and robust controllable dialogue generation.

Limitations

Our method improves control performance while maintaining sentence fluency by leveraging entropy-based control strength, and it enables flu-

ent sentence generation in both single- and multi-attribute settings. However, this approach has several limitations. First, the dataset used in this study is limited in both quantity and diversity, and validation was performed only on dialogues with two attributes. Therefore, construction of new corpus with additional attributes is necessary to evaluate the generalizability of the proposed method. Second, in multi-attribute settings, a refined normalization process must be developed to account for the number of classes for each attribute as the number of attribute classifiers increases. To address these limitations, future work should construct more diverse datasets and explore effective probability control strategies under multi-attribute conditions.

Ethics Statement

The proposed method aims to enhance the interest and accuracy of responses generated by chatbots to improve user experience. However, this method could be potentially used for malicious purposes. In our experiments, we focus on attributes like emotion and dialog-act, but if malicious desired attributes such as bias are used, the model could be induced to generate inappropriate responses. Therefore, generating controlled responses using malicious attributes should be restricted.

References

Kushal Arora, Kurt Shuster, Sainbayar Sukhbaatar, and Jason Weston. 2022. Director: Generator-classifiers for supervised language modeling. *Preprint*, arXiv:2206.07694.

PF Budzianowski and Iñigo Casanueva. 2018. Towards end-to-end multi-domain dialogue modelling.

Jiwei Li, Michel Galley, Chris Brockett, Jianfeng Gao, and Bill Dolan. 2016a. A diversity-promoting objective function for neural conversation models. *Preprint*, arXiv:1510.03055.

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*.

Yanran Li, Hui Su, Xiaoyu Shen, Wenjie Li, Ziqiang Cao, and Shuzi Niu. 2017. DailyDialog: A manually labelled multi-turn dialogue dataset. In *Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 986–995, Taipei, Taiwan. Asian Federation of Natural Language Processing.

Chin-Yew Lin. 2004. Rouge: A package for automatic evaluation of summaries. In *Text summarization branches out*, pages 74–81.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized bert pretraining approach. *Preprint*, arXiv:1907.11692.

John Morris, Eli Lifland, Jin Yong Yoo, Jake Grigsby, Di Jin, and Yanjun Qi. 2020. Textattack: A framework for adversarial attacks, data augmentation, and adversarial training in nlp. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 119–126.

Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever, et al. 2019. Language models are unsupervised multitask learners. *OpenAI blog*, 1(8):9.

Claude Elwood Shannon. 2001. A mathematical theory of communication. *ACM SIGMOBILE mobile computing and communications review*, 5(1):3–55.

Hugo Touvron, Louis Martin, Kevin Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov, Soumya Batra, Prajjwal Bhargava, Shruti Bhosale, et al. 2023. Llama 2: Open foundation and fine-tuned chat models. *arXiv preprint arXiv:2307.09288*.

Alex Warstadt, Amanpreet Singh, and Samuel R. Bowman. 2019. Neural network acceptability judgments. *Preprint*, arXiv:1805.12471.

Kevin Yang and Dan Klein. 2021. Fudge: Controlled text generation with future discriminators. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Association for Computational Linguistics.

Weihao Zeng, Lulu Zhao, Keqing He, Ruotong Geng, Jingang Wang, Wei Wu, and Weiran Xu. 2023. Seen to unseen: Exploring compositional generalization of multi-attribute controllable dialogue generation. *Preprint*, arXiv:2306.10317.

Yizhe Zhang, Siqi Sun, Michel Galley, Yen-Chun Chen, Chris Brockett, Xiang Gao, Jianfeng Gao, Jingjing Liu, and Bill Dolan. 2020. Dialogpt: Large-scale generative pre-training for conversational response generation. *Preprint*, arXiv:1911.00536.

Zhiling Zhang, Mengyue Wu, and Kenny Q. Zhu. 2023. Semantic space grounded weighted decoding for multi-attribute controllable dialogue generation. *Preprint*, arXiv:2305.02820.

A Analysis

A.1 POS-Based Entropy Analysis

To investigate the theoretical foundation of entropy-aware control strength, we conducted a token-level entropy analysis based on part-of-speech (POS) tags. Entropy scores are measured using the token probability distributions from the DialoGPT models.

In Table 7, we can observe some patterns in the relationship between a part-of-speech tag and its entropy. Functional words such as ADP (adpositions), DET (determiners), and PUNCT (punctuation) tend to have relatively low entropy because they are typically governed by strict grammatical constraints that limit lexical variability during generation. In contrast, content words including NOUN (nouns), VERB (verbs), ADJ (adjectives), and ADV (adverbs) exhibit higher entropy because these words often occur in more flexible contexts and are responsible for conveying core semantic information, which results in a broader distribution over candidate tokens.

These findings provide a theoretical support for our entropy-based control strategy. Specifically, positions with high entropy exhibit greater lexical variability, which allows for more flexible attribute manipulation without compromising the fluency of the generated text. In contrast, low-entropy tokens are typically constrained by grammatical structure and should therefore be subjected to weaker control in order to maintain grammatical correctness. This POS-tag-based analysis confirms that entropy can serve as an effective indicator for determining the appropriate timing and location of control application during the generation process.

POS Tag	DialoGPT-S	DialoGPT-L
ADP	3.04	3.19
CCONJ	3.68	4.04
PUNCT	3.69	3.75
PART	3.78	3.51
SCONJ	4.20	4.37
DET	4.27	4.26
PRON	4.35	4.40
ADV	4.41	4.55
NUM	4.51	4.71
X	4.55	4.94
VERB	4.58	4.74
NOUN	4.81	5.39
SYM	4.85	5.31
PROPN	4.85	5.29
ADJ	5.00	5.35
AUX	5.10	4.93
INTJ	5.47	5.71

Table 7: Entropy measurement score in each part-of-speech (POS) tag. Results are reported for the DialoGPT models: S (small) and L (large). Items in Table 7 are sorted in ascending order based on the entropy scores of DialoGPT-S.

A.2 Control Strength Analysis

In this study, to mathematically model control strength as inversely proportional to the entropy of the token distribution, we define control strength as a function of entropy and employ three representative decreasing functions. These functions are designed to dynamically adjust control strength depending on the level of entropy, thereby enabling more assertive control in confident contexts (i.e., low entropy) and weaker control when the model is uncertain (i.e., high entropy).

The three control weighting functions are as follows:

$$exponetial = 1 + exp(-attr_entropy) \quad (11)$$

$$negative = 1 + (log(V) - attr_entropy) \quad (12)$$

$$reciprocal = 1 + (\frac{1}{attr_entropy}) \quad (13)$$

When comparing the three functions, the linear function maintains a direct linear relationship between entropy and weight. In contrast, both the exponential function and the inverse function yield similar weights in high-entropy regions, but diverge significantly in low-entropy regions with assigning notably larger weights.

Table 8 shows that the linear function tends to degrade performance in multi-attribute control tasks. This is due to the fact that its weights decrease linearly without convergence as entropy increases, resulting in unstable control. On the other hand, both the exponential and reciprocal functions exhibit convergence as entropy increases, and outperform the baseline model. These two functions produce larger differences in weights in low-entropy regions, and our experiments confirm that the reciprocal function particularly achieves strong performance. This is more likely because, in cases where

Model	Accuracy(emo)	Accuracy(act)	Rouge-1	Rouge-L	Dist-1	Dist-2	Grammar
DialoGPT	-	-	9.00	8.53	0.58	0.76	90.21
Director	80.48	60.65	9.41	8.99	0.58	0.73	90.22
- `lm_entropy`	79.24	58.63	9.49	9.04	0.58	0.73	**90.31**
- `exponential`	80.48	61.12	9.42	9.01	0.58	0.73	90.21
- `negative`	80.17	60.96	9.41	9.00	0.58	0.73	90.21
- `reciprocal`	**81.18**	**61.20**	9.49	8.97	0.58	0.73	90.23

Table 8: Experiments on control strength application methods. We compare the performance of the lm_entropy method, which utilizes the probability distribution of the language model, with three alternative methods that use the probability distribution of the attribute classifier.

the probability distribution is highly concentrated (i.e., one token is assigned a very high probability), the reciprocal function is better suited to apply strong control.

Furthermore, the reciprocal function provides more advantages than the exponential function by offering a broader range of control strength in low-entropy settings. Consequently, we conclude that the reciprocal weighting function is well-suited for assigning control strength appropriately under both low-entropy and high-entropy conditions.

B Cases of ECO decoding

Dialogue history

A : Let's program your courses. Since you have had English literature, you should take American literature and American prose and fiction. You transcript indicate that your English background is strong, so I don't think you have any problem with it.

B : How many credits for each course?

A : Three, you also should take two three-credit-course in education department.

Response generation

Desired attribute : **Commissive**

Director

B : three credits... You're gonna have to take two thirds of course. . .

+ ECO decoding

B : Thanks, I'll take that course.

Figure 5: In case where the response fails to satisfy the desired attribute with the existing method but satisfies the desired attribute using ECO decoding.

Dialogue history

A : I want something sweet after dinner.

B : What do you have in mind?

Response generation

Desired attribute : **Inform**

Director

A : I'm not sure what I want in return.

+ ECO decoding

A : I'm going to eat a whole box of cereal.

Figure 6: In case where it fails to generate a context-consistent response with the existing method but generates a context-consistent response using ECO decoding.

C Licenses

The DailyDialog dataset is licensed under CC BY-NC-SA 4.0 License. The DialoGPT model is licensed under Contributor License Agreement (CLA) and Llama2 model is licensed under Meta Llama 2 Community License Agreement. The RoBERTa-based CoLA grammaticality model is licenced under MIT License.