Adaptive Planning for Multi-Attribute Controllable Summarization with Monte Carlo Tree Search

Anonymous ACL submission

Abstract

Controllable summarization moves beyond generic outputs toward human-aligned summaries guided by specified attributes. In practice, the interdependence among attributes makes it challenging for language models to satisfy correlated constraints consistently. Moreover, previous approaches often require per-attribute fine-tuning, limiting flexibility across diverse summary attributes. In this paper, we propose adaptive planning for multi-attribute controllable summarization (PACO), a training-free framework that reframes the task as planning the order of sequential attribute control with a customized Monte Carlo Tree Search (MCTS). In PACO, nodes represent summaries, and actions correspond to single-attribute adjustments, enabling progressive refinement of only the attributes requiring further control. This strategy adaptively discovers optimal control orders, ultimately producing summaries that effectively meet all constraints. Extensive experiments across diverse domains and models demonstrate that PACO achieves robust multi-attribute controllability, surpassing both LLM-based self-planning models and fine-tuned baselines. Remarkably, PACO with Llama-3.2-1B rivals the controllability of the much larger Llama-3.3-70B baselines. With larger models, PACO achieves superior control performance, outperforming all competitors.

1 Introduction

Controllable summarization, which tailors summaries to user-specified attributes such as *length*, *extractiveness*, or *topic*, is essential for real-world applications, enabling more personalized outputs. For instance, a student preparing for an exam may prefer a concise summary highlighting only the key topics, whereas a teacher preparing lecture materials may require a more detailed version with high specificity with broad coverage.

Recent works have explored multi-attribute controllable summarization training with attribute-specific supervision. For instance, Goyal et al. (2022) leveraged a mixture-of-experts (MoE) with each decoder specializing in one attribute, while Zhang et al. (2023) employed hard prompt tuning (HP) and soft prefix tuning (SP) to train a model for multiple attributes. However, these methods require additional fine-tuning for each attribute, limiting flexibility and generalization to unseen preferences. More fundamentally, the autoregressive generation of language models may struggle to enforce multiple correlated constraints simultaneously in a single decoding pass (Figure 1).

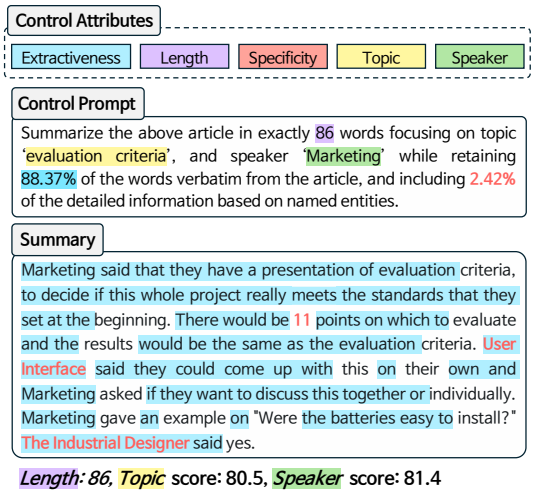

Figure 1: Summaries consist of multiple attributes. Our goal is to generate outputs that satisfy diverse user-specified constraints simultaneously.

As multiple attributes often interact in complex ways, achieving full control of all attributes might structurally lead to conflicts; for example, improving extractiveness may inadvertently compromise length control. Moreover, the space of possible attribute control orders grows combinatorially, leaving the unsolved question of how to systematically explore the optimal and effective control paths.

To address these challenges, we propose adaptive planning for multi-attribute controllable summarization (PACO), a training-free framework that

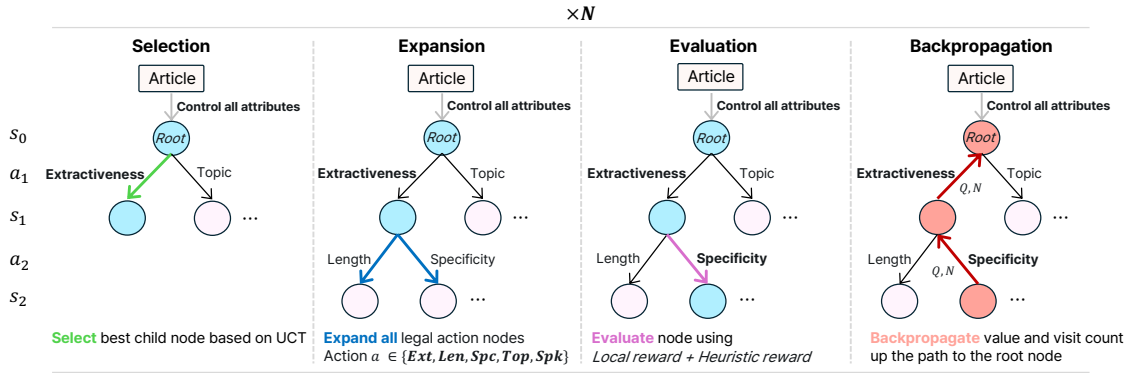

Figure 2: Illustration of the MCTS process in PACO. The tree search begins from a summary generated with a prompt that requests control over all attributes, serving as the root node. After all simulations are completed, the node with the highest *degree* is selected from the entire tree during the decision stage.

transforms multi-attribute summarization into a sequential decision-making (i.e., planning) problem. Instead of attempting to enforce all constraints at once, PACO progressively adjusts attributes step by step. Specifically, we design a tailored Monte Carlo Tree Search (MCTS) algorithm that explores different control orders at each step by defining nodes at the summary level, while allowing revisiting attributes to adaptively find an optimal control path. Since each node encapsulates a complete summary, we can select the summary that maximizes the degree of attribute control once the tree is fully expanded. To ensure the structured search and evaluation, we categorize attributes by type, distinguishing between *deterministic* attributes, which must match exact user targets, and *non-deterministic* attributes where higher values are preferable.

We evaluate PACO on diverse domains, including $\text{MACSum}_{\text{Doc}}$, $\text{MACSum}_{\text{Dial}}$ (Zhang et al., 2023), and DialogSum (Chen et al., 2021). In experiments with a range of LLMs, PACO demonstrates robust control performance across models with different sizes and domains. Remarkably, our training-free PACO with a 1B model achieves control performance comparable to that of a 70B baseline model, and PACO with a 70B model outperforms all baselines across all attributes, showing consistently strong controllability. Crucially, PACO achieves these controllability gains without sacrificing summary quality by incrementally adjusting attributes rather than forcing all constraints at once, which may risk compromising summary quality. Our main contributions are as follows:

- We introduce PACO, the first framework to cast controllable summarization as a sequential planning problem and adapt MCTS to systematically explore optimal control paths.
- We define summary-level nodes and categorize attributes by type to assign rewards, thereby enabling flexible and effective enforcement of multiple attribute constraints.
- Extensive experiments across models and datasets demonstrate that PACO achieves superior controllability and strong alignment with user preferences, even without attribute-specific training.

2 LLM-Based Self-Planning

Since LLMs struggle to control multiple attributes simultaneously (Ryu et al., 2025), we aim to adjust attributes progressively, one at a time. However, due to interdependence among the attributes, determining an effective control order remains challenging. To examine whether LLMs can perform attribute control planning on their own, we introduce two prompt-based self-planning baselines: implicit and explicit self-planning.

Implicit self-planning. We prompt the LLMs to generate the summary in a single pass while implicitly considering which attribute to control first, using *Let's think step-by-step* (Kojima et al., 2022). The model is encouraged to consider the control order without explicitly generating a separate plan, and to reflect this consideration in the output.

Explicit self-planning. We refer to prompting the model to control all attributes at once as the initial summary baseline. Based on this initial summary, which may not satisfy all constraints, we prompt the model to generate a control plan specifying the order of adjustments. This explicit plan

indicates the sequence in which to modify the misaligned attributes in the initial summary. Guided by the plan, the model sequentially adjusts each attribute, generating intermediate summaries at each step. Once all attributes in the plan have been controlled, the final summary serves as the output. We introduce two explicit self-planning strategies: a *base* version, where the LLM plans the order for attributes control without imposing any constraints or guidance; and an *adaptive* version, which is weakly guided to flexibly select and order only the attributes deemed necessary, potentially revisiting the same attribute multiple times depending on the context. Refer to Appendix G for detailed self-planning prompts.

3 PACO

Controlling multiple attributes is nontrivial as the outcome depends on the order of control and the search space of possible orders is combinatorially large. In addition, control attempts may succeed immediately or fail repeatedly, making fixed strategies unreliable and motivating the need for systematic exploration. To optimize attribute control planning, we propose PACO, which integrates the MCTS algorithm into multi-attribute controllable summarization. We formulate the attribute control planning process as a Markov Decision Process (MDP). Defining nodes at a fine granularity, such as the token or sentence level, as in prior tree-based approaches with LLMs (Yao et al., 2023; Hao et al., 2023; Wan et al., 2024), can lead to an intractably large search space especially in tasks that require long-form generation, such as text summarization. To address this, we define each node at the summary level, reducing search complexity and easing the planning burden on the model.

LLMs serve as the policy π, where each action a corresponds to controlling a single attribute. We start from an initial summary that reflects all attribute controls, which serves as the root node s_0, and adaptively search for the optimal attribute control order $[attribute_1, attribute_2, ..., attribute_n]$ for each document. Each intermediate summary serves as a state s, forming a sequence of transitions from s_0 through successive attribute adjustments. At each step t, the model determines the action a_t to control a specific attribute and generates the next summary s_{t+1} by taking the full history $s_0, s_1, ..., s_t$ as input. This allows it to make informed decisions based on all preceding modifications. We define the maximum tree width w as the number of legal actions and denote the tree depth as d. The algorithm iterates until it reaches the terminal state T, which occurs when all attributes have been successfully controlled or when the step limit is exceeded. Figure 2 presents an overview of PACO, and Figure 3 provides an example of how each attribute is adjusted. Starting from the initial summary, PACO identifies under-controlled attributes and progressively adjusts them according to the planned control order, ultimately producing a summary aligned with the target values.

Selection. The PACO process begins at the root node s_0, which is generated by prompting the model to control all attributes in a single initial attempt. The algorithm then explores the search tree by selecting nodes based on a variant of the Predictor Upper Confidence Tree (PUCT) (Rosin, 2011) algorithm, using the following equation:

$$U(s,a) = c_{\text{puct}} \cdot \pi_\theta(s,a) \cdot \frac{\sqrt{\sum_b N(s,b)}}{1 + N(s,a)} \quad (1)$$

$$a = \arg\max_a \left[Q(s,a) + U(s,a)\right] \quad (2)$$

Here, $Q(s,a)$ denotes the state-action value and $N(s,a)$ is the visit count for action a at state s, both of which are maintained and updated during the search. $N(s,b)$ denotes the visit count of the action b taken from state s, where b is one of the possible actions at that state. To balance exploration and exploitation, we use the following term $c_{\text{puct}} = \log\left(\frac{\sum_b N(s,b)+c_{\text{base}}+1}{c_{\text{base}}}\right) + c_{\text{init}}$, which encourages exploration of less-visited actions while promoting the exploitation of those with high value estimates to maximize expected reward. The selection process continues until a terminal state (T) is reached, defined either as a summary that satisfies all attribute constraints or upon reaching a predefined maximum tree depth.

Expansion. When a leaf node is reached, we expand it by generating child nodes for all possible actions. The action space is defined as $action \in \{ext, len, spc, top, spk\}$, as each action corresponds to controlling a single attribute. Since the effect of a previously applied action can be altered by subsequent actions, all actions are considered legal throughout the search process.

Evaluation. To estimate the value of a node, we introduce two rewards: a local reward based on intermediate steps, which captures immediate

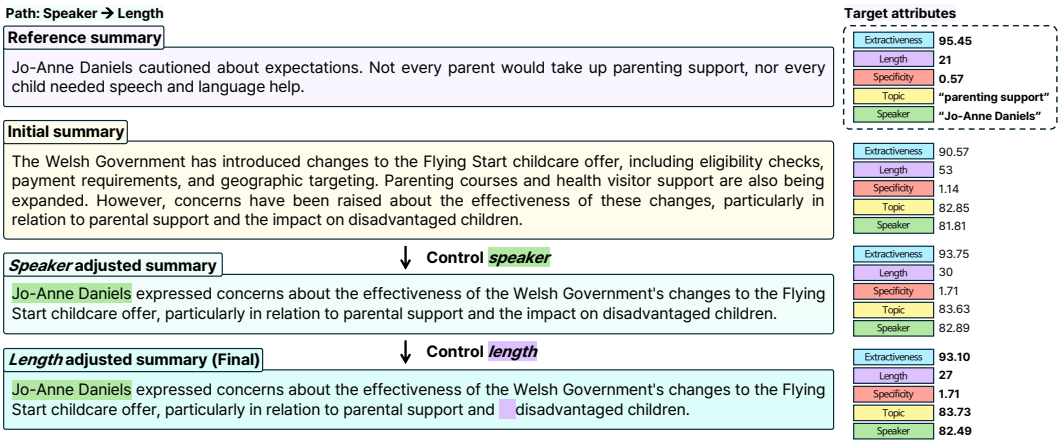

Figure 3: An example of PACO adjusting a summary through its planning process. The initial summary shows that LLMs struggle with multiple attribute constraints in a single pass. To address this, PACO successfully refines the summary to meet target attributes. indicates shifts to speaker-focused content; highlights removal of unnecessary details to reach the target length. Values beside the reference summary indicate the target attributes, while values beside generated summaries show their measured attribute scores.

improvements, and a heuristic score that reflects global confidence in the final output. The local reward is computed by adapting multi-attribute measurements for controllable summarization (Ryu et al., 2025), with each attribute defined as follows:

- *Extractiveness*: the proportion of summary words that appear in the source document.
- *Length*: the total word count of the summary.
- *Specificity*: the ratio of named entities to the total number of words in the summary.
- *Topic*: the average embedding similarity $\mathcal{B}$ between n summary words and k topic words: $\frac{1}{k}\sum_{j\in k}\frac{1}{n}\sum_{i\in s}\mathcal{B}(\text{topic}_j, \text{word}_i)$.
- *Speaker*: the embedding similarity between the summary and a set of utterances $\mathcal{U}$ from a target speaker in the dialogue, measured by $\text{BERTScore}(s, \mathcal{U})$.

Using these attribute measurements, we compute the mean absolute deviation (MAD) between the predicted and target values for each requested attribute. We distinguish *deterministic* and *non-deterministic* attributes: *deterministic* attributes, such as *extractiveness*, *length*, and *specificity*, are expected to match user-specified target values, whereas *non-deterministic* attributes, such as *topic* and *speaker*, are evaluated based on their alignment with the target, where higher values indicate better alignments. Therefore, we use the alignment score itself for *non-deterministic* attributes instead of MAD. The total local reward, referred to as the control *degree*, is computed by averaging MAD over *deterministic* attributes (avg_{det}) and adding alignment scores for *non-deterministic* attributes ($avg_{\text{non-det}}$). Since a lower value of avg_{det} indicates better performance, we take its reciprocal to align the reward direction. These hyperparameters can be adjusted to control the relative importance of *deterministic* and *non-deterministic* attributes.

$$\text{Local reward} = \frac{\alpha}{avg_{\text{det}} + \varepsilon} + \frac{1}{\beta} \cdot avg_{\text{non-det}} \quad (3)$$

While traditional MCTS approaches utilize rollouts to estimate the value function (Kocsis and Szepesvári, 2006; Gelly and Silver, 2011), MCTS applications in LLMs typically employ prompt-based heuristic value functions (Yao et al., 2023; Hao et al., 2023; Yu et al., 2023) or learned value functions (Wan et al., 2024; Chen et al., 2024) to reduce computational cost. In a similar manner, we design a heuristic value function tailored to the controllable summarization. Specifically, we define a heuristic score that assesses whether the model can feasibly control all remaining attributes given the current summary and the action path taken so far. Since it is difficult for the model to generate this score as an explicit numeric value, we frame the query as a binary question, and use the probability of a "`Yes`" response as the heuristic score.

$$\text{Heuristic score} = p(\texttt{Yes} \mid s) \quad (4)$$

Backpropagation. At the end of each simulation, we update the visit count and cumulative value

estimate $W(s, a)$ for each node along the search path using the simulation result $V(s_l)$ from the leaf node s_l. The mean action-value $Q(s, a)$ is then obtained as the ratio of the cumulative value to the visit count.

$$N(s_t, a_t) \leftarrow N(s_t, a_t) + 1 \quad (5)$$

$$W(s_t, a_t) \leftarrow W(s_t, a_t) + V(s_L) \quad (6)$$

$$Q(s_t, a_t) = \frac{W(s_t, a_t)}{N(s_t, a_t)} \quad (7)$$

Decision. While node exploration during simulation is guided by stepwise value updates, the final summary is selected based on a fixed *degree*. Unlike standard MCTS approaches that select the most visited or highest-value leaf node (Browne et al., 2012), PACO selects the node with the highest degree across the entire tree. This enables PACO to adaptively control a subset of attributes, rather than enforce all of them, allowing for more flexible summarization tailored to each document. See Appendix A for algorithmic details.

4 Experimental Setup

Datasets. We conduct experiments on two mixed-attribute controllable summarization datasets, $\text{MACSum}_{\text{Dial}}$ and $\text{MACSum}_{\text{Doc}}$ (Zhang et al., 2023), and on a topic-focused dialogue summarization dataset, DialogSum (Chen et al., 2021). $\text{MACSum}_{\text{Dial}}$ is constructed from the QMSum dataset (Zhong et al., 2021), which contains meeting transcripts from three sources: AMI (Carletta et al., 2005), ICSI (Janin et al., 2003), and committee meetings from the Welsh Parliament and the Parliament of Canada. $\text{MACSum}_{\text{Doc}}$ is based on the CNN/DailyMail (See et al., 2017), a news domain dataset. DialogSum consists of real-life scenario that covers general topics from everyday life. Notably, only $\text{MACSum}_{\text{Dial}}$ includes the *speaker*.

Models. We demonstrate the robustness of our approach by applying it to various LLMs of different sizes, including the Llama series (Llama-3.2-1B-Instruct and Llama-3.3-70B-Instruct) (Touvron et al., 2023; Grattafiori et al., 2024) and Qwen2.5-7B-Instruct (Bai et al., 2023; Yang et al., 2024). As baselines, we compare against LLM-based self-planning methods, namely implicit self-planning and explicit self-planning (including both *base* and *adaptive* versions), as well as hard prompt tuning combined with soft prefix tuning (HP+SP) (Raffel et al., 2020; Li and Liang, 2021), reimplemented following Zhang et al. (2023) based on $\text{BART}_{\text{large}}$ (Lewis et al., 2020). We measure embedding similarity using BERTScore (Zhang et al., 2020) and extract named entities with FLAIR (Akbik et al., 2019), a well-established named entity recognition (NER) model trained on OntoNotes 5 (Pradhan et al., 2013), which covers various domains, including news and conversational speech.

Metrics. We adopt different evaluation for *deterministic* and *non-deterministic* attributes. While we compute the mean absolute deviation (MAD) between the target and generated attribute values for *deterministic* attributes (lower is better), we directly evaluate the generated values for *non-deterministic* attributes (higher is better). Although our main goal is to achieve controllable summarization, maintaining overall summary quality remains important. To this end, we additionally evaluate the quality of the generated summaries using ROUGE-1 (Lin, 2004) and BERTScore F1 (Zhang et al., 2020).

5 Main Results

Controllability results. While target attributes can be arbitrarily chosen, we use the values of the reference summaries to enable direct comparison. For *topic* and *speaker*, we use values provided in the dataset. In our main results, we use only the local reward as the node value. As shown in Table 1, smaller-scale LLM baselines struggle to control attributes, particularly *length*, resulting in excessively high MAD on $\text{MACSum}_{\text{Dial}}$. Since this dataset contains long and complex meeting transcripts, even large models such as Llama-3.3-70B struggle to control *length*, with MAD exceeding 15. In contrast, PACO consistently shows strong control performance across different models. Notably, it reduces the MAD for *length* from 55.68 to 17.96 on the 1B model, which is comparable to the performance of the 70B baseline. On Llama-3.3-70B, PACO achieves an average MAD of approximately 5 over *deterministic* attributes, demonstrating precise control and clearly outperforming all baselines. Applying PACO to Qwen2.5-7B also yields substantial performance gains over the base model. Its control ability falls between those of the 1B and 70B Llama models, further highlighting both effectiveness and generalizability.

Robustness across datasets. As shown in Table 2, PACO again significantly outperforms all baselines on $\text{MACSum}_{\text{Doc}}$. Remarkably, the 1B PACO model even surpasses the 70B baseline,

Model	# of Params	Ext (↓)	Len (↓)	Spc (↓)	Top (↑)	Spk (↑)	ROUGE (↑)	BERTScore (↑)
Reference summary	-	0.00	0.00	0.00	0.796	0.802	-	-
HP+SP ($BART_{large}$)*	406M	6.66	34.66	7.08	0.807	**0.804**	0.315	0.871
Llama 3.2 Instruct	1B	10.79	55.68	9.30	0.783	0.795	0.270	0.854
Qwen 2.5 Instruct	7B	9.70	17.82	6.99	0.797	0.795	0.301	0.867
Llama 3.3 Instruct	70B	6.43	15.72	7.11	0.800	0.798	0.328	0.871
Implicit self-planning	70B	7.35	27.70	8.09	0.802	0.795	0.304	0.869
Explicit self-planning	70B	7.44	28.19	7.32	**0.808**	0.794	0.287	0.869
Explicit self-planning+	70B	7.08	24.52	7.32	0.801	0.795	0.312	0.869
PACO (Llama 3.2 Instruct)	1B	9.30	17.96	7.22	0.792	0.794	0.288	0.859
PACO (Qwen 2.5 Instruct)	7B	8.72	11.79	5.43	0.799	0.794	0.302	0.868
PACO (Llama 3.3 Instruct)	70B	**4.91**	**7.63**	**3.81**	0.795	0.798	0.328	0.869

Table 1: Controllability evaluation results on $MACSum_{Dial}$. Explicit self-planning denotes the *base* version, while explicit self-planning+ refers to the *adaptive* variant. Bold indicates the best controllability within the dataset, and * marks models trained on the corresponding data.

Model	# of Params	Ext (↓)	Len (↓)	Spc (↓)	Top (↑)	ROUGE (↑)	BERTScore (↑)
Reference summary	-	0.00	0.00	0.00	0.806	-	-
HP+SP ($BART_{large}$)*	406M	9.04	19.43	4.55	0.803	0.327	0.881
Llama 3.2 Instruct	1B	7.84	9.67	4.27	0.792	0.330	0.879
Qwen 2.5 Instruct	7B	7.39	13.78	5.42	0.793	0.321	0.879
Llama 3.3 Instruct	70B	8.03	10.91	3.58	0.802	0.308	0.878
Implicit self-planning	70B	8.14	17.05	4.56	0.803	0.296	0.875
Explicit self-planning	70B	8.22	15.64	3.89	**0.808**	0.285	0.875
Explicit self-planning+	70B	8.19	13.86	3.99	0.800	0.286	0.873
PACO (Llama 3.2 Instruct)	1B	7.53	**4.58**	3.73	0.796	0.326	0.879
PACO (Qwen 2.5 Instruct)	7B	6.37	7.03	4.18	0.797	0.319	0.880
PACO (Llama 3.3 Instruct)	70B	**4.49**	4.81	**2.84**	0.794	0.322	0.876

Table 2: Controllability evaluation results on $MACSum_{Doc}$, which does not include the *speaker* attribute.

Model	Params	Ext (↓)	Len (↓)	Spc (↓)	Top (↑)
Reference summary	-	0.00	0.00	0.00	0.817
Llama 3.2 Instruct	1B	20.45	15.65	52.43	0.815
Qwen 2.5 Instruct	7B	12.08	5.20	26.62	0.817
Llama 3.3 Instruct	70B	14.91	2.26	20.82	**0.829**
PACO (Llama 3.2 Instruct)	1B	14.17	6.28	28.48	0.825
PACO (Qwen 2.5 Instruct)	7B	8.71	3.30	19.14	0.820
PACO (Llama 3.3 Instruct)	70B	**8.35**	**1.56**	**10.20**	0.828

Table 3: Evaluation results on DialogSum. While annotator-specific attributes lead to varying control trends, PACO consistently outperforms all baselines.

while our 70B model exhibits dominant controllability, clearly surpassing all other models. Compared to $MACSum_{Dial}$, which consists of longer and more complex input texts, all models shows better controllability on $MACSum_{Doc}$ with simpler inputs. These results highlight that PACO maintains robust control across domains and input complexities, whereas baseline methods show a notable decline as inputs become longer and more complex.

We further evaluate PACO on DialogSum. Table 3 shows that PACO achieves substantial gains in controllability across all model sizes. Interestingly, the controllability pattern on DialogSum differs from that of the MACSum datasets. While *length* is the most difficult and *specificity* the easiest to control in MACSum, the reverse holds for DialogSum. This discrepancy may stem from domain-specific properties or variation in annotation style, as human-written summaries can differ across annotators. These results underscore the effectiveness of adaptive control in PACO, which flexibly adjusts to the unique characteristics of each dataset.

Balancing between attribute types. The results show that LLM-based models are more effective at controlling *non-deterministic* than *deterministic* attributes, with *topic* and *speaker* scores comparable to the reference summaries. Given that attribute types can be prioritized, we place higher weight on the *deterministic* attributes. We provide more detailed experiments on balancing control performance across attribute types in Appendix D. Although HP+SP is explicitly trained for attribute control, it often fails to follow instructions. We consider this to be due to structural constraints of the encoder–decoder architecture and to the use of vague supervision for *deterministic* attributes (e.g.,"high" rather than precise targets).

Comparison with self-planning. We evaluate whether LLMs can perform attribute control planning (Tables 1, 2). The results show that both implicit and explicit self-planning fail to generate effective plans, performing even worse than the baseline. In particular, implicit self-planning exhibits

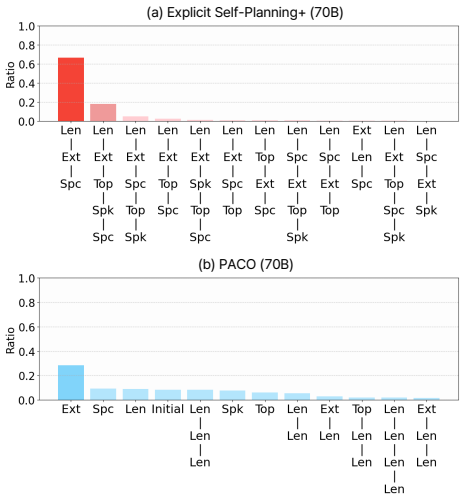

Figure 4: (a) LLMs often over-control and lack diversity in their self-generated plans, whereas (b) PACO controls only the necessary attributes for each instance. We visualize the top 10 plans for each method.

the weakest control performance. The *adaptive* version, explicit self-planning+, incorporates soft constraints into the prompt and improves on the base version; however, it still lags behind the baseline. These findings highlight that LLMs struggle with attribute planning in multi-attribute controllable, underscoring the need for a more effective planning strategy to guide the generation process.

As shown in Figure 4, PACO selectively adjusts only the necessary attributes starting from the initial summary, resulting in diverse and well-balanced control plan distributions. In contrast, explicit self-planning+, despite being prompted to plan only for necessary adjustments, tends to produce repetitive and imbalanced plans across most data points. This highlights that LLMs struggle with planning in controllable summarization.

Quality evaluation. Focusing too heavily on attribute control may risk degrading the summary quality, so we also evaluate the overall summary quality (Table 1, 2). Notably, by incrementally controlling attributes rather than enforcing all constraints simultaneously, PACO avoids potential quality degradation and preserves summary quality comparable to the baseline. Although LLMs have already demonstrated strong summarization capabilities (Goyal et al., 2023; Pu et al., 2023; Zhang et al., 2024b; Ryu et al., 2024b), PACO not only excels in control performance but also preserves their high generation quality. Additionally, although LLMs tend to generate more paraphrased outputs, which often lead to lower ROUGE scores

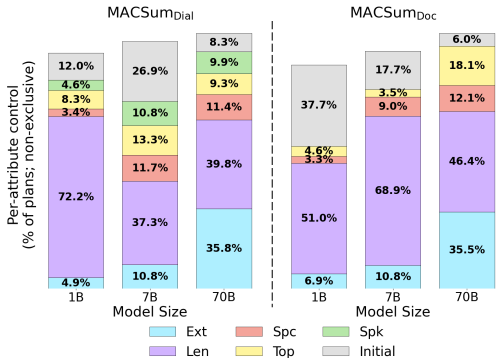

Figure 5: Each bar shows the frequency of attribute control per model size, with repeated attributes in a plan counted once. Percentages indicate relative proportions. *Initial* refers to the state where all attributes were controlled simultaneously at the beginning.

compared to trained encoder-decoder models, they can still achieve higher ROUGE scores when given precise control instructions.

6 Analysis

Frequency of attribute control. In Figure 5, we analyze the attributes adjusted by PACO across model sizes and domains. As the model size increases, fewer initial summary is selected. This suggests that larger models adjust attributes more effectively. Particularly, they more often control *extractiveness* and *specificity*, demonstrating their capacity for sophisticated control. Across all model sizes, *length* is the most frequently adjusted attribute, likely due to its strong correlation with others and the resulting need for additional revisions. In MACSum$_{Dial}$, which has longer and more complex inputs, the initial summary is selected less in total than the MACSum$_{Doc}$. This indicates that longer inputs often require additional adjustments to satisfy multiple constraints.

Step-wise control patterns. We analyze the average attribute controlled by PACO at each step across variable-length plans in DialogSum (Figure 6). 70B model adjusts diverse attributes in the early steps, but the later steps tend to focus on *length* control, as it can be highly affected by other attributes. In contrast, the 1B model shows no controllability gains with deeper adjustments. In particular, it struggled to control attributes beyond length, rarely attempting such adjustments.

Ablation study on the value function. We present an ablation study comparing different strategies for computing node values, comparing control degree at the current step with heuristic scoring, a

Model	# of Params	Extractiveness (↓)	Length (↓)	Specificity (↓)	Topic (↑)	Speaker (↑)
PACO (L)	1B	**9.30**	17.96	**7.22**	0.792	**0.794**
PACO (H)	1B	9.56	21.98	7.97	0.791	0.793
PACO (L + H)	1B	9.56	**16.12**	7.63	**0.793**	**0.794**
PACO (L)	70B	**4.91**	7.63	**3.81**	0.795	**0.798**
PACO (H)	70B	5.06	7.59	3.94	**0.796**	**0.798**
PACO (L + H)	70B	5.16	**7.56**	4.28	0.795	0.796

Table 4: Ablation study of the value function. 'L' denotes the local reward, and 'H' denotes the heuristic score.

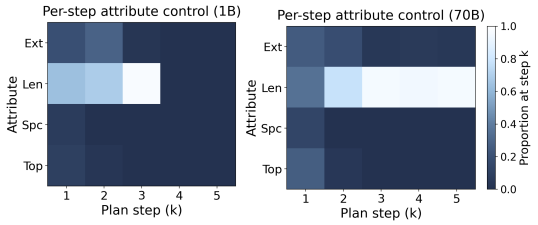

Figure 6: Average attribute controlled at each step. Later steps mostly control *length*, which may be due to high correlation with other attributes.

common choice in LLM-based MCTS (Table 4). The results show that the heuristic score provides little benefit as a value function, especially for the 1B model. Combining both signals offers only marginal gains, which are insufficient to offset the additional cost. Thus, using only the local reward is both more efficient and effective, likely because predicting whether a partially controlled summary will meet all remaining attributes is nontrivial.

7 Related Work

Controllable summarization. Prior work on controllable summarization has focused primarily on single-attribute control (Zhong et al., 2021; Liu and Chen, 2021; Dou et al., 2021; Mao et al., 2022; Zhang et al., 2022; Bahrainian et al., 2022; Ahuja et al., 2022; Liu et al., 2022; Maddela et al., 2022; Mehra et al., 2023; Xu et al., 2023; Pagnoni et al., 2023; Wang et al., 2023; Chan et al., 2021; Ryu et al., 2025), most commonly targeting adjustments to *length* and *topic* (Urlana et al., 2024). To control diverse attributes, He et al. (2022) introduced a framework capable of controlling various attributes, such as *entity*, *length*, and *purpose*, though these were not controlled simultaneously.

Recently, there has been growing interest in controlling multiple attributes simultaneously (Fan et al., 2018; Goyal et al., 2022; Zhang et al., 2023). Zhang et al. (2023) introduced mixed-attribute controllable summarization, while Goyal et al. (2022) leveraged MoE to control multiple attributes jointly. However, these methods require additional training for each attribute, making them impractical when the number of attributes increases. In contrast, PACO leverages LLMs without any attribute-specific training, using planning via MCTS to discover optimal control paths and enable simultaneous control of all target attributes.

Tree search for LLMs. Tree search has predominantly been applied to reasoning tasks, where problems are decomposed into sub-questions and represented as nodes in a search tree to facilitate step-by-step reasoning toward the correct answer (Yao et al., 2023; Hao et al., 2023; Wan et al., 2024; Chen et al., 2024; Zhang et al., 2024a; Xie et al., 2024; Lee et al., 2025). Yao et al. (2023) frame each node as a partial solution and search the tree to solve complex problems. Hao et al. (2023) treats the language model as a world model, defining task-specific states and actions. Whereas prior work has applied MCTS to identify reasoning paths that lead to a correct final answer, our task requires that the entire decoding process satisfy multiple constraints. To address this, we tailor MCTS to the controllable summarization setting, defining each node at the summary level rather than at the token or sentence level. In addition, since the degree of control over a previously adjusted attribute may change as subsequent actions are taken, we allow the same attribute to be adjusted multiple times during the search.

8 Conclusion

We propose PACO, an adaptive planning method that incorporates Monte Carlo Tree Search into multi-attribute controllable summarization, enabling effective control over multiple attributes. Since it is challenging for language models to enforce all constraints simultaneously in a single pass, PACO incrementally adjusts attributes by constructing effective control paths. Unlike LLM-based self-planning methods, PACO revises only the necessary attributes. As a result, it demonstrates robust and consistent control across various models and domains while preserving summary quality.

Limitations

Although PACO offers strong controllability across multiple attributes, several limitations remain. First, while summary-level nodes help reduce the search space, the tree search remains computationally expensive (refer to Appendix E). Finding an optimal control path requires deeper simulations, leading to longer runtimes and limited scalability. To address this, future work could explore search-time heuristics to reduce computational cost without sacrificing control quality. Second, incorporating the optimization of quality dimensions, as explored in previous work such as Ryu et al. (2024a) and Song et al. (2025), could expand controllability beyond attribute alignment to broader quality dimensions such as *coherence*, *consistency*, *relevance*, and *fluency*. To support more comprehensive and user-tailored summarization, future work could extend PACO to accommodate a broader range of attribute types.

Ethical Statement

This paper focuses on applications in controllable summarization and raises no ethical concerns. All datasets used are publicly available, and AI assistance was employed exclusively for grammar correction.

References

Ojas Ahuja, Jiacheng Xu, Akshay Gupta, Kevin Horecka, and Greg Durrett. 2022. ASPECTNEWS: Aspect-oriented summarization of news documents. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 6494–6506, Dublin, Ireland. Association for Computational Linguistics.

Alan Akbik, Tanja Bergmann, Duncan Blythe, Kashif Rasul, Stefan Schweter, and Roland Vollgraf. 2019. FLAIR: An easy-to-use framework for state-of-the-art NLP. In *NAACL 2019, 2019 Annual Conference of the North American Chapter of the Association for Computational Linguistics (Demonstrations)*, pages 54–59.

Seyed Ali Bahrainian, Sheridan Feucht, and Carsten Eickhoff. 2022. NEWTS: A corpus for news topic-focused summarization. In *Findings of the Association for Computational Linguistics: ACL 2022*, pages 493–503, Dublin, Ireland. Association for Computational Linguistics.

Jinze Bai et al. 2023. Qwen technical report. *Preprint*, arXiv:2309.16609.

Cameron B. Browne, Edward Powley, Daniel Whitehouse, Simon M. Lucas, Peter I. Cowling, Philipp Rohlfshagen, Stephen Tavener, Diego Perez, Spyridon Samothrakis, and Simon Colton. 2012. A survey of monte carlo tree search methods. *IEEE Transactions on Computational Intelligence and AI in Games*, 4(1):1–43.

Jean Carletta, Simone Ashby, Sebastien Bourban, Mike Flynn, Mael Guillemot, Thomas Hain, Jaroslav Kadlec, Vasilis Karaiskos, Wessel Kraaij, Melissa Kronenthal, Guillaume Lathoud, Mike Lincoln, Agnes Lisowska, Iain McCowan, Wilfried Post, Dennis Reidsma, and Pierre Wellner. 2005. The ami meeting corpus: a pre-announcement. In *Proceedings of the Second International Conference on Machine Learning for Multimodal Interaction*, MLMI'05, page 28–39, Berlin, Heidelberg. Springer-Verlag.

Hou Pong Chan, Lu Wang, and Irwin King. 2021. Controllable summarization with constrained Markov decision process. *Transactions of the Association for Computational Linguistics*, 9:1213–1232.

Guoxin Chen, Minpeng Liao, Chengxi Li, and Kai Fan. 2024. Alphamath almost zero: Process supervision without process. In *Advances in Neural Information Processing Systems*, volume 37, pages 27689–27724. Curran Associates, Inc.

Yulong Chen, Yang Liu, Liang Chen, and Yue Zhang. 2021. DialogSum: A real-life scenario dialogue summarization dataset. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 5062–5074, Online. Association for Computational Linguistics.

Zi-Yi Dou, Pengfei Liu, Hiroaki Hayashi, Zhengbao Jiang, and Graham Neubig. 2021. GSum: A general framework for guided neural abstractive summarization. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 4830–4842, Online. Association for Computational Linguistics.

Angela Fan, David Grangier, and Michael Auli. 2018. Controllable abstractive summarization. In *Proceedings of the 2nd Workshop on Neural Machine Translation and Generation*, pages 45–54, Melbourne, Australia. Association for Computational Linguistics.

Sylvain Gelly and David Silver. 2011. Monte-carlo tree search and rapid action value estimation in computer go. *Artif. Intell.*, 175:1856–1875.

Tanya Goyal, Nazneen Rajani, Wenhao Liu, and Wojciech Kryscinski. 2022. HydraSum: Disentangling style features in text summarization with multi-decoder models. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 464–479, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.

Tanya Goyal et al. 2023. News summarization and evaluation in the era of gpt-3. *Preprint*, arXiv:2209.12356.

Aaron Grattafiori et al. 2024. The llama 3 herd of models. *Preprint*, arXiv:2407.21783.

Shibo Hao, Yi Gu, Haodi Ma, Joshua Hong, Zhen Wang, Daisy Wang, and Zhiting Hu. 2023. Reasoning with language model is planning with world model. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 8154–8173, Singapore. Association for Computational Linguistics.

Junxian He, Wojciech Kryscinski, Bryan McCann, Nazneen Rajani, and Caiming Xiong. 2022. CTRLsum: Towards generic controllable text summarization. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 5879–5915, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.

A. Janin, D. Baron, J. Edwards, D. Ellis, D. Gelbart, N. Morgan, B. Peskin, T. Pfau, E. Shriberg, A. Stolcke, and C. Wooters. 2003. The icsi meeting corpus. In *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03).*, volume 1, pages I–I.

Levente Kocsis and Csaba Szepesvári. 2006. Bandit based monte-carlo planning. In *Proceedings of the 17th European Conference on Machine Learning*, ECML'06, page 282–293, Berlin, Heidelberg. Springer-Verlag.

Takeshi Kojima, Shixiang Shane Gu, Machel Reid, Yutaka Matsuo, and Yusuke Iwasawa. 2022. Large language models are zero-shot reasoners. In *Proceedings of the 36th International Conference on Neural Information Processing Systems*, NIPS '22, Red Hook, NY, USA. Curran Associates Inc.

Sungjae Lee, Hyejin Park, Jaechang Kim, and Jungseul Ok. 2025. Semantic exploration with adaptive gating for efficient problem solving with language models. *Preprint*, arXiv:2501.05752.

Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin Stoyanov, and Luke Zettlemoyer. 2020. BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 7871–7880, Online. Association for Computational Linguistics.

Xiang Lisa Li and Percy Liang. 2021. Prefix-tuning: Optimizing continuous prompts for generation. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 4582–4597, Online. Association for Computational Linguistics.

Chin-Yew Lin. 2004. ROUGE: A package for automatic evaluation of summaries. In *Text Summarization Branches Out*, pages 74–81, Barcelona, Spain. Association for Computational Linguistics.

Yizhu Liu, Qi Jia, and Kenny Zhu. 2022. Length control in abstractive summarization by pretraining information selection. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 6885–6895, Dublin, Ireland. Association for Computational Linguistics.

Zhengyuan Liu and Nancy Chen. 2021. Controllable neural dialogue summarization with personal named entity planning. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 92–106, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.

Mounica Maddela, Mayank Kulkarni, and Daniel Preotiuc-Pietro. 2022. EntSUM: A data set for entity-centric extractive summarization. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 3355–3366, Dublin, Ireland. Association for Computational Linguistics.

Ziming Mao, Chen Henry Wu, Ansong Ni, Yusen Zhang, Rui Zhang, Tao Yu, Budhaditya Deb, Chenguang Zhu, Ahmed Awadallah, and Dragomir Radev. 2022. DYLE: Dynamic latent extraction for abstractive long-input summarization. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1687–1698, Dublin, Ireland. Association for Computational Linguistics.

Dhruv Mehra, Lingjue Xie, Ella Hofmann-Coyle, Mayank Kulkarni, and Daniel Preotiuc-Pietro. 2023. EntSUMv2: Dataset, models and evaluation for more abstractive entity-centric summarization. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 5538–5547, Singapore. Association for Computational Linguistics.

Artidoro Pagnoni, Alex Fabbri, Wojciech Kryscinski, and Chien-Sheng Wu. 2023. Socratic pretraining: Question-driven pretraining for controllable summarization. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 12737–12755, Toronto, Canada. Association for Computational Linguistics.

Sameer Pradhan, Alessandro Moschitti, Nianwen Xue, Hwee Tou Ng, Anders Björkelund, Olga Uryupina, Yuchen Zhang, and Zhi Zhong. 2013. Towards robust linguistic analysis using OntoNotes. In *Proceedings of the Seventeenth Conference on Computational Natural Language Learning*, pages 143–152, Sofia, Bulgaria. Association for Computational Linguistics.

Xiao Pu, Mingqi Gao, and Xiaojun Wan. 2023. Summarization is (almost) dead. *arXiv preprint arXiv:2309.09558*.

Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J. Liu. 2020. Exploring the limits of transfer learning with a unified text-to-text transformer. *J. Mach. Learn. Res.*, 21(1).

Christopher D. Rosin. 2011. Multi-armed bandits with episode context. *Annals of Mathematics and Artificial Intelligence*, 61(3):203–230.

Sangwon Ryu, Heejin Do, Daehee Kim, Hwanjo Yu, Dongwoo Kim, Yunsu Kim, Gary Geunbae Lee, and Jungseul Ok. 2025. Exploring iterative controllable summarization with large language models. *Preprint*, arXiv:2411.12460.

Sangwon Ryu, Heejin Do, Yunsu Kim, Gary Lee, and Jungseul Ok. 2024a. Multi-dimensional optimization for text summarization via reinforcement learning. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 5858–5871, Bangkok, Thailand. Association for Computational Linguistics.

Sangwon Ryu, Heejin Do, Yunsu Kim, Gary Geunbae Lee, and Jungseul Ok. 2024b. Key-element-informed sllm tuning for document summarization. In *Interspeech 2024*, pages 1940–1944.

Julian Schrittwieser, Ioannis Antonoglou, Thomas Hubert, Karen Simonyan, Laurent Sifre, Simon Schmitt, Arthur Guez, Edward Lockhart, Demis Hassabis, Thore Graepel, Timothy Lillicrap, and David Silver. 2020. Mastering atari, go, chess and shogi by planning with a learned model. *Nature*, 588(7839):604–609.

Abigail See, Peter J. Liu, and Christopher D. Manning. 2017. Get to the point: Summarization with pointer-generator networks. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1073–1083, Vancouver, Canada. Association for Computational Linguistics.

David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, Adrian Bolton, Yutian Chen, Timothy Lillicrap, Fan Hui, Laurent Sifre, George van den Driessche, Thore Graepel, and Demis Hassabis. 2017. Mastering the game of go without human knowledge. *Nature*, 550(7676):354–359.

Hwanjun Song, Taewon Yun, Yuho Lee, Jihwan Oh, Gihun Lee, Jason Cai, and Hang Su. 2025. Learning to summarize from LLM-generated feedback. In *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pages 835–857, Albuquerque, New Mexico. Association for Computational Linguistics.

Hugo Touvron et al. 2023. Llama: Open and efficient foundation language models. *Preprint*, arXiv:2302.13971.

Ashok Urlana, Pruthwik Mishra, Tathagato Roy, and Rahul Mishra. 2024. Controllable text summarization: Unraveling challenges, approaches, and prospects - a survey. In *Findings of the Association for Computational Linguistics: ACL 2024*, pages 1603–1623, Bangkok, Thailand. Association for Computational Linguistics.

Ziyu Wan, Xidong Feng, Muning Wen, Stephen Marcus Mcaleer, Ying Wen, Weinan Zhang, and Jun Wang. 2024. AlphaZero-like tree-search can guide large language model decoding and training. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pages 49890–49920. PMLR.

Bin Wang, Zhengyuan Liu, and Nancy Chen. 2023. Instructive dialogue summarization with query aggregations. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 7630–7653, Singapore. Association for Computational Linguistics.

Yuxi Xie, Anirudh Goyal, Wenyue Zheng, Min-Yen Kan, Timothy P. Lillicrap, Kenji Kawaguchi, and Michael Shieh. 2024. Monte carlo tree search boosts reasoning via iterative preference learning. *Preprint*, arXiv:2405.00451.

Ruochen Xu, Song Wang, Yang Liu, Shuohang Wang, Yichong Xu, Dan Iter, Pengcheng He, Chenguang Zhu, and Michael Zeng. 2023. LMGQS: A large-scale dataset for query-focused summarization. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 14764–14776, Singapore. Association for Computational Linguistics.

An Yang et al. 2024. Qwen2.5 technical report. *arXiv preprint arXiv:2412.15115*.

Shunyu Yao, Dian Yu, Jeffrey Zhao, Izhak Shafran, Tom Griffiths, Yuan Cao, and Karthik Narasimhan. 2023. Tree of thoughts: Deliberate problem solving with large language models. In *Advances in Neural Information Processing Systems*, volume 36, pages 11809–11822. Curran Associates, Inc.

Xiao Yu, Maximillian Chen, and Zhou Yu. 2023. Prompt-based Monte-Carlo tree search for goal-oriented dialogue policy planning. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 7101–7125, Singapore. Association for Computational Linguistics.

Di Zhang, Xiaoshui Huang, Dongzhan Zhou, Yuqiang Li, and Wanli Ouyang. 2024a. Accessing gpt-4 level mathematical olympiad solutions via monte carlo tree self-refine with llama-3 8b. *Preprint*, arXiv:2406.07394.

Tianyi Zhang, Varsha Kishore, Felix Wu, Kilian Q. Weinberger, and Yoav Artzi. 2020. Bertscore: Evaluating text generation with bert. In *Proceedings of the International Conference on Learning Representations*.

Tianyi Zhang, Faisal Ladhak, Esin Durmus, Percy Liang, Kathleen McKeown, and Tatsunori B. Hashimoto. 2024b. Benchmarking large language models for news summarization. *Transactions of the Association for Computational Linguistics*, 12:39–57.

Yusen Zhang, Yang Liu, Ziyi Yang, Yuwei Fang, Yulong Chen, Dragomir Radev, Chenguang Zhu, Michael Zeng, and Rui Zhang. 2023. MACSum: Controllable summarization with mixed attributes. *Transactions of the Association for Computational Linguistics*, 11:787–803.

Yusen Zhang, Ansong Ni, Ziming Mao, Chen Henry Wu, Chenguang Zhu, Budhaditya Deb, Ahmed Awadallah, Dragomir Radev, and Rui Zhang. 2022. Summ^n: A multi-stage summarization framework for long input dialogues and documents. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1592–1604, Dublin, Ireland. Association for Computational Linguistics.

Ming Zhong, Da Yin, Tao Yu, Ahmad Zaidi, Mutethia Mutuma, Rahul Jha, Ahmed Hassan Awadallah, Asli Celikyilmaz, Yang Liu, Xipeng Qiu, and Dragomir Radev. 2021. QMSum: A new benchmark for query-based multi-domain meeting summarization. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 5905–5921, Online. Association for Computational Linguistics.

A PACO Algorithm

Algorithm 1 outlines the PACO procedure. The algorithm includes selection, expansion, evaluation, and backpropagation during the simulation phase, followed by a decision step that selects the final summary from the entire tree.

B Hyperparameters

We set the maximum tree depth d to 5 and perform 8 simulations per search. To compute the local reward, we use α = 1 for *deterministic* attributes and β = 10 for *non-deterministic* attributes, balancing their scales. We adopt most MCTS hyperparameters from Silver et al. (2017) and Schrittwieser et al. (2020), including c_{base} = 19652 and c_{init} = 1.25. we follow Zhang et al. (2023) to train the HP+SP baseline and use the final checkpoint.

C Hardware Usage

We used 4 NVIDIA A100-SXM4-80GB GPUs for our experiments.

D Balancing Between Deterministic and Non-deterministic Attributes

To evaluate each node during the search, we compute the local reward as $\frac{\alpha}{avg_{\text{det}}+\varepsilon} + \frac{1}{\beta} \cdot avg_{\text{non-det}}$, where avg_{det} denotes the average deviation from the target values for *deterministic* attributes, and $\text{avg}_{\text{non-det}}$ represents the average affinity score for *non-deterministic* attributes. The hyperparameters α and β control the relative importance of the two terms, and ε is a small constant added to ensure numerical stability. Since LLMs generally struggle more with structural, *deterministic* attributes (e.g., *length*, *extractiveness*) than with content-related, *non-deterministic* attributes, we set $\beta = 10$ in the main results (Tables 1 and 2) to mildly upweight *deterministic* control performance.

In Table 5, we vary the value of β to adjust the relative weighting between *deterministic* and *non-deterministic* attributes in our experiments. We conduct experiments on the $\text{MACSum}_{\text{Dial}}$ dataset using Qwen2.5-7B and Llama-3.3-70B as baseline models. A smaller β value increases the relative weight assigned to *non-deterministic* attributes. The experimental results show that as β decreases, the scores for *non-deterministic* attributes such as *topic* and *speaker* gradually increase, while those for *deterministic* attributes like *extractiveness*, *length*, and *specificity* tend to decrease. Specifically, with the

Algorithm 1 PACO ($\mathcal{M}_\theta$)

Require: LM $\mathcal{M}_\theta$, attribute measure f
Require: article x, target attributes a^*
Require: controllable attributes $\mathcal{A}$, hyperparameters: simulations n, max depth d

// initialize root node with summary controlling all attributes
$y_0 \leftarrow \mathcal{M}_\theta(x, a^*, \mathcal{A})$
$\hat{a}_0 \leftarrow f(x, y_0)$
$deg_0 \leftarrow \text{degree}(\hat{a}_0, a^*)$
$s_0 \leftarrow \text{node}(y_0, \hat{a}_0, deg_0, \text{depth} = 0)$
for $i = 1$ to n **do**
  initialize state $s \leftarrow s_0$
  // selection
  while s is not a leaf and not terminated **and** $\text{depth}(s) < d$ **do**
    $a \leftarrow \arg\max_{a' \in \mathcal{A}} \text{PUCT}(s, a')$
    $s \leftarrow \text{child}(s, a)$
  end while
  // expansion
  if s is a leaf and not terminated **then**
    create child nodes $\{\text{child}(s, a')\}_{a' \in \mathcal{A}}$
  end if
  // evaluation
  $a' \leftarrow \arg\max_{a' \in \mathcal{A}} \text{PUCT}(s, a')$
  $s' \leftarrow \text{child}(s, a')$
  if s' has no summary **then**
    $y \leftarrow \mathcal{M}_\theta(x, a^*, a', \text{history}(s'))$
    $\hat{a} \leftarrow f(x, y)$
    $deg \leftarrow \text{degree}(\hat{a}, a^*)$
    store y, $\hat{a}$, deg in s'
  else
    retrieve y, $\hat{a}$, deg from s'
  end if
  // backpropagation
  while $s' \neq s_0$ **do**
    update stats at s'
    $s' \leftarrow$ parent of s'
  end while
end for
// decision
$s^* \leftarrow \arg\max_{s \in \text{Tree}} \text{degree}(s)$
return summary y^* from s^*

Model	# of Params	β	Extractiveness (↓)	Length (↓)	Specificity (↓)	Topic (↑)	Speaker (↑)
Reference summary	-	-	0.00	0.00	0.00	0.796	0.802
Qwen 2.5 Instruct	7B	-	9.70	17.82	6.99	0.797	0.795
PACO (Qwen 2.5 Instruct)	7B	10	**8.72**	**11.79**	**5.43**	0.799	0.794
PACO (Qwen 2.5 Instruct)	7B	0.5	8.98	12.66	5.96	0.799	0.794
PACO (Qwen 2.5 Instruct)	7B	0.2	8.85	13.22	5.99	0.800	0.794
PACO (Qwen 2.5 Instruct)	7B	0.1	8.88	13.45	5.99	0.800	0.794
PACO (Qwen 2.5 Instruct)	7B	0.01	9.31	17.94	6.44	**0.801**	**0.795**
Llama 3.3 Instruct	70B	-	6.43	15.72	7.11	0.800	0.798
PACO (Llama 3.3 Instruct)	70B	10	4.91	**7.63**	**3.81**	0.795	0.798
PACO (Llama 3.3 Instruct)	70B	0.5	**4.83**	8.48	4.34	0.796	0.799
PACO (Llama 3.3 Instruct)	70B	0.2	4.90	11.01	4.81	0.798	0.800
PACO (Llama 3.3 Instruct)	70B	0.1	5.04	12.41	5.39	0.799	0.800
PACO (Llama 3.3 Instruct)	70B	0.01	5.88	15.77	5.80	**0.800**	**0.801**

Table 5: Effect of weighting attribute types. Adjusting the weights shifts the emphasis between *deterministic* and *non-deterministic* attributes. As β decreases, more weight is placed on *non-deterministic* attributes, resulting in improved scores for *topic* and *speaker*.

Model	# of Params	Time (s) per Summary
Baseline	70B	22.826
Implicit self-planning	70B	33.410
Explicit self-planning	70B	104.312
Explicit self-planning+	70B	90.941
PACO	70B	196.380

Table 6: Average time required to generate a single summary.

Qwen2.5-7B when $\beta = 10$, the MAD for length was 11.79, and the topic score was 0.799. However, when β was reduced to 0.01 to emphasize *non-deterministic* attributes, the MAD for length increased to 17.94, while the topic score improved to 0.801. A similar trend was observed with the Llama-3.3-70B model. These findings suggest that users can control the summarization output by adjusting weights to emphasize the attributes they value most.

E Computational Costs

In Table 6, we present the average time required by each model to generate a final summary. Although PACO incurs a higher computational cost, it clearly outperforms self-planning approaches, which are also relatively expensive. This demonstrates a favorable trade-off between computation and controllability, particularly in tasks that require structured control. Importantly, our method tackles the complex challenge of multi-attribute control entirely through test-time inference. Since higher computational cost is often necessary for stronger reasoning and controllability, and LMs are becoming faster and more efficient, we consider the increased computational cost required for stronger controllability to be practical and promising rather than a long-term limitation.

F Attribute Control Prompts

The following are the detailed prompts used for attribute control. These prompts are shared across both PACO and the self-planning methods to ensure a fair comparison.

F.1 Initial prompts

```
You are a helpful assistant. Your
task is to generate adjusted
summary for user.

article:
{{Article}}

Summarize the above article
in exactly {{length}} words
focusing on {{topic}} and
{{speaker}} while retaining
{{extractiveness}} of the words
verbatim from the article, and
including {{specificity}} of
the detailed information based
on named entities. Ensure
the summary is well-written,
logically sound, with clear
sentence flow.

summary (generate summary ONLY):
```

F.2 Extractiveness control prompts

```
You are a helpful assistant. Your
```

task is to generate adjusted summary for user.

article:
{{Article}}

{{History}}

summary:
{{Previous summary}}

The summary you previously generated did not follow the given instructions. Summarize the above article while retaining {{extractiveness}} of the words verbatim from the article. Ensure the summary is well-written, logically sound, with clear sentence flow.

summary (generate summary ONLY):

F.3 Length control prompts

You are a helpful assistant. Your task is to generate adjusted summary for user.

article:
{{Article}}

{{History}}

summary:
{{Previous summary}}

The summary you previously generated did not follow the given instructions. Summarize the above article in exactly {{length}} words. Ensure the summary is well-written, logically sound, with clear sentence flow.

summary (generate summary ONLY):

F.4 Specificity control prompts

You are a helpful assistant. Your task is to generate adjusted summary for user.

article:
{{Article}}

{{History}}

summary:
{{Previous summary}}

The summary you previously generated did not follow the given instructions. Summarize the above article including {{specificity}} of the detailed information based on named entities. Ensure the summary is well-written, logically sound, with clear sentence flow.

summary (generate summary ONLY):

F.5 Topic control prompts

You are a helpful assistant. Your task is to generate adjusted summary for user.

article:
{{Article}}

{{History}}

summary:
{{Previous summary}}

The summary you previously generated did not follow the given instructions. Summarize the above article focusing on topic {{topic}}. Ensure the summary is well-written, logically sound, with clear sentence flow.

summary (generate summary ONLY):

F.6 Speaker control prompts

You are a helpful assistant. Your task is to generate adjusted summary for user.

article:
{{Article}}

{{History}}

summary:
{{Previous summary}}

The summary you previously generated did not follow the given instructions. Summarize the above article focusing on speaker {{speaker}}. Ensure the summary is well-written, logically sound, with clear sentence flow.

summary (generate summary ONLY):

G LLM-Based Self-Planning

G.1 Implicit self-planning

You are a helpful assistant. Your task is to generate adjusted summary for user.

article:
{{Article}}

{{Initial prompts}}

summary:
{{Previous summary}}

You must modify {{target attributes}}, but since it is difficult to modify them all at once, you should adjust them one by one. Let's think step by step, internally consider the order in which to adjust the attributes, and gradually revise the summary to generate one that satisfies all attributes.

summary (generate summary ONLY):

G.2 Explicit self-planning

You are a helpful assistant. Your task is to generate adjusted summary for user.

article:
{{Article}}

{{Initial prompts}}

summary:
{{Previous summary}}

You are a helpful assistant. Your task is to generate a plan for adjusting summary.

You must modify {{target attributes}}, but since it is difficult to modify them all at once, you should adjust them one by one. Plan which attribute should be modified first. The output should be returned as a list. For example, plan = ['attribute1', 'attribute2', ...]

plan (generate plan ONLY):

G.3 Explicit self-planning+

You are a helpful assistant. Your task is to generate adjusted summary for user.

article:
{{Article}}

{{Initial prompts}}

summary:
{{Previous summary}}

You are a helpful assistant. Your task is to generate a plan for adjusting summary.

You must modify {{target attributes}}, but since it is difficult to modify them all at once, you should adjust them one by one. Plan which attribute should be modified first. Note that you do not need to modify all attributes, and you may adjust the same attribute multiple times if necessary. The output should be returned as a list. For example, plan = ['attribute1', 'attribute2',

```
...]

plan (generate plan ONLY):
```

H LLM-based Heuristic Score

```
You are a helpful assistant. Your task is to assess the summary for user.

article:
{{Article}}

summary:
{{Summary}}

Can further adjustments to this summary fully satisfy all target attributes? The final objective is to generate a summary that meets the target attributes {{target attributes}}. The current summary has been adjusted in the order of {{path}}. Keep in mind that changes made to earlier attributes might be disrupted when you adjust later ones.

answer (generate Yes or No ONLY):
```