Offre n°2025-08818

PhD Position F/M Optimizing Controllable Text Generation: A Comparative Study of Alignment Strategies and Inference-Time Scaling

Le descriptif de l'offre ci-dessous est en Anglais

Type de contrat : CDD

Niveau de diplôme exigé : Bac + 5 ou équivalent

Fonction : Doctorant

Contexte et atouts du poste

la thèse se déroulera dans le cadre de PR[AI]RIE-PSAI et sera co-encadrée par Marc LELARGE et Guillaume BAUDART.

Non-discrimination, ouverture et transparence. L'ensemble des partenaires de PR[AI]RIE-PSAI s'engagent à soutenir et promouvoir l'égalité, la diversité et l'inclusion au sein de ses communautés. Nous encourageons les candidatures issues de profils variés, que nous veillerons à sélectionner via un processus de recrutement ouvert et transparent

Mission confiée

Large language models (LLMs) have demonstrated remarkable capabilities in generating coherent and contextually relevant text across a wide range of domains, including open-ended dialogue, code synthesis, and formal reasoning. However, steering these models to produce outputs that align with human-defined goals, domain constraints, and task-specific requirements remains a persistent challenge. This PhD project seeks to investigate and compare several complementary approaches to controllable text generation, focusing on three primary families of techniques: Supervised Fine-Tuning (SFT), Reinforcement Learning (RL), and

Controlled Decoding.

In addition to these established methods, the project will explore an increasingly recognized yet underexplored dimension of model optimization: inference-time compute scaling. Recent studies suggest that dedicating more computational resources or adopting more sophisticated generation algorithms at inference time — including token-level selection, meta-generation, and efficient reranking — can significantly improve output quality without altering the base model parameters.

The proposed research will combine theoretical analysis and experimental evaluation to map the trade-offs between alignment accuracy, computational efficiency, and generalization capabilities across different strategies. Special attention will be given to structured, rule-based tasks such as code generation and formal theorem proving, where control, correctness, and logical consistency are especially critical. The expected outcomes include reproducible benchmarks, new hybrid alignment methods, and the development of inference-time optimization techniques that can improve the quality and reliability of LLM outputs across diverse applications.

Principales activités

The overarching goal of this PhD project is to deepen the scientific understanding of controllable text generation by systematically evaluating and enhancing model alignment techniques for large language models (LLMs). This research will focus on three principal alignment families — Supervised Fine-Tuning (SFT), Reinforcement Learning (RL), and Controlled Decoding — and will extend this analysis to include inference-time scaling strategies, which have shown increasing promise in recent research. The project will particularly emphasize structured and rule-governed domains such as **code generation** and **formal theorem proving**.

1. Comparative Study of Alignment Techniques

Supervised Fine-Tuning (SFT): Investigate SFT as a baseline alignment strategy, where labeled datasets are used to condition the model toward desired behaviors. This approach will be analyzed for its ability to enforce task-specific accuracy and domain adherence.

Reinforcement Learning (RL): Explore reinforcement learning techniques, including KL-regularized RL and Reinforcement Learning with Human Feedback (RLHF), as methods for fine-tuning LLMs in settings where explicit labels are scarce but reward signals can guide alignment toward human-defined objectives.

Controlled Decoding: Study controlled decoding strategies such as prefix scoring, blockwise decoding, and token-level filtering to steer output at inference time without modifying the model’s underlying weights. This line of inquiry focuses on low-overhead control and real-time adaptability.

2. Incorporation of Inference-Time Compute Scaling

This research will also focus on the emerging role of inference-time compute scaling as a means to improve the quality of model outputs without retraining or modifying the underlying model parameters. Beyond the well-known benefits of scaling compute during training, recent work has highlighted that more sophisticated inference-time strategies — including token-level generation algorithms, meta-generation techniques that rerank multiple candidate outputs, and efficiency-oriented decoding frameworks — can significantly enhance alignment and output diversity. The project will explore how these inference-time approaches contribute to controllability and how they can be combined with Supervised Fine-Tuning, Reinforcement Learning, and Controlled Decoding to strike a balance between computational efficiency and output quality, especially in scenarios where training-time resources are limited or inference is constrained by real-world application demands.

We will design and experiment with hybrid pipelines that combine SFT, RL, Controlled Decoding, and inference-time scaling techniques to create alignment strategies that balance control, flexibility, and computational cost.

3. Application to Structured Generation Tasks

The effectiveness of the proposed alignment and inference-time scaling strategies will be evaluated through a combination of domain-specific and general-purpose metrics. For structured generation tasks such as code synthesis and formal theorem proving, particular attention will be paid to the syntactic and structural correctness of the outputs, ensuring they conform to the expected formal languages and formats. Logical and semantic coherence will be assessed to verify that the generated content is not only grammatically correct but also factually and deductively sound. In the case of theorem proving, proof verification success rates will be measured using automated proof checkers to ensure formal validity. Finally, the efficiency of generation will be systematically analyzed by considering both the computational cost of each method and the quality of the outputs, highlighting the trade-offs between resource usage and alignment performance.

4. Expected Contributions

- A comprehensive and reproducible comparative analysis of SFT, RL, and Controlled Decoding in structured text generation settings.
- Development of novel hybrid approaches for combining alignment techniques effectively.
- Open-source tools and benchmarks for assessing controllability in code generation and formal reasoning.

- Potential peer-reviewed publications and dissemination of research findings at leading machine learning and natural language processing venues.

References:

- Keskar, N. S., McCann, B., Varshney, L. R., Xiong, C., & Socher, R. (2019). CTRL: A Conditional Transformer Language Model for Controllable Generation. *arXiv preprint arXiv:1909.05858*

- Jones, A.L. (2021) Scaling Scaling Laws with Board Games. *arxiv preprint arXiv:2104.03113*

- Mudgal, S., Lee, J., Ganapathy, H., Li, Y., Wang, T., Huang, Y., & Beirami, A. (2024). Controlled Decoding from Language Models. *arXiv preprint arXiv:2310.17022*

- Willard B.T., Louf R. (2023) Efficient Guided Generation for Large Language Models. *arXiv preprint arXiv:2307.09702*

- Liang, X., Wang, H., Wang, Y., Song, S., Yang, J., Niu, S., Hu, J., Liu, D., Yao, S., Xiong, F., & Li, Z. (2024). Controllable Text Generation for Large Language Models: A Survey. *arXiv preprint arXiv:2408.12599*

- Welleck S., Bertsch A., Finlayson M., Schoelkopf H., Xie A., Neubig G., Kulikov I., Harchaoui Z. (2024). From Decoding to Meta-Generation: Inference-time Algorithms for Large Language Models . *arXiv preprint arXiv:2406.16838*

Compétences

Skills and Tools

- **Programming:** Python, PyTorch
- **Machine Learning:** NLP, Transformer-based models, Reinforcement Learning
- **Formal verification:** Rocq, Lean.
- **Data Processing:** Hugging Face Transformers

Avantages

- Subsidized meals
- Partial reimbursement of public transport costs
- Leave: 7 weeks of annual leave + 10 extra days off due to RTT (statutory reduction in working hours) + possibility of exceptional leave (sick children, moving home, etc.)
- Possibility of teleworking
- Flexible organization of working hours
- Professional equipment available (videoconferencing, loan of computer equipment, etc.)
- Social, cultural and sports events and activities
- Access to vocational training
- Social security coverage

Informations générales

- **Thème/Domaine :** Optimisation, apprentissage et méthodes statistiques Statistiques (Big data) (BAP E)
- **Ville :** Paris
- **Centre Inria :** Centre Inria de Paris
- **Date de prise de fonction souhaitée :** 2025-09-01
- **Durée de contrat :** 3 ans
- **Date limite pour postuler :** 2025-05-23

Contacts

- **Équipe Inria :** ARGO
- **Directeur de thèse :**
  Lelarge Marc / Marc.Lelarge@inria.fr

A propos d'Inria

Inria est l’institut national de recherche dédié aux sciences et technologies du numérique. Il emploie 2600 personnes. Ses 215 équipes-projets agiles, en général communes avec des partenaires académiques, impliquent plus de 3900 scientifiques pour relever les défis du numérique, souvent à l’interface d’autres disciplines. L’institut fait appel à de nombreux talents dans plus d’une quarantaine de métiers différents. 900 personnels d’appui à la recherche et à l’innovation contribuent à faire émerger et grandir des projets scientifiques ou entrepreneuriaux qui impactent le monde. Inria travaille avec de nombreuses entreprises et a accompagné la création de plus de 200 start-up. L'institut s'e?orce ainsi de répondre aux enjeux de la transformation numérique de la science, de la société et de l'économie.

Attention: Les candidatures doivent être déposées en ligne sur le site Inria. Le traitement des candidatures adressées par d'autres canaux n'est pas garanti.

Consignes pour postuler

Required documents :
- a resume;
- a one-page cover letter describing the applicant's ambitions for the subject described and the relevance of the application to the subject description;
- copies of most recent diplomas.

Sécurité défense :
Ce poste est susceptible d'être affecté dans une zone à régime restrictif (ZRR), telle que définie dans le décret n°2011-1425 relatif à la protection du potentiel scientifique et technique de la nation (PPST). L'autorisation d'accès à une zone est délivrée par le chef d'établissement, après avis ministériel favorable, tel que défini dans l'arrêté du 03 juillet 2012, relatif à la PPST. Un avis ministériel défavorable pour un poste affecté dans une ZRR aurait pour conséquence l'annulation du recrutement.

Politique de recrutement :
Dans le cadre de sa politique diversité, tous les postes Inria sont accessibles aux personnes en situation de handicap.