Can Textual Gradient Work in Federated Learning?

Author: Minghui Chen[1,2,3], Ruinan Jin[1,2], Wenlong Deng[1,2], Yuanyuan Chen[3], Zhi Huang[4], Han Yu[3], Xiaoxiao Li[1,2]

Affiliation: [1]The University of British Columbia, [2]Vector Institute, [3]Nanyan Technological University, [4]University of Pennsylvania

Speaker: Minghui Chen
Date: 2025/4/1

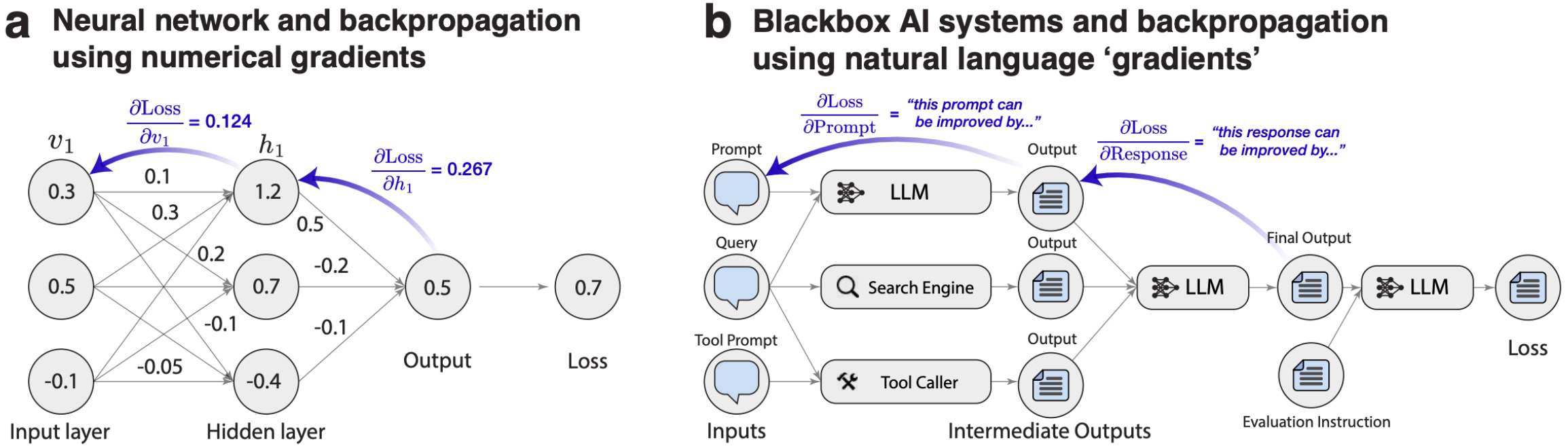
ICLR
International Conference On Learning Representations

FedTextGrad

0. Overview: TextGrad Background

TextGrad: An automated framework that **optimizes compound AI systems** by **backpropagating** LLM feedback through natural language suggestions. It improves tasks like question answering, molecule design, and treatment planning **without manual tuning**, following PyTorch-like syntax for flexibility.

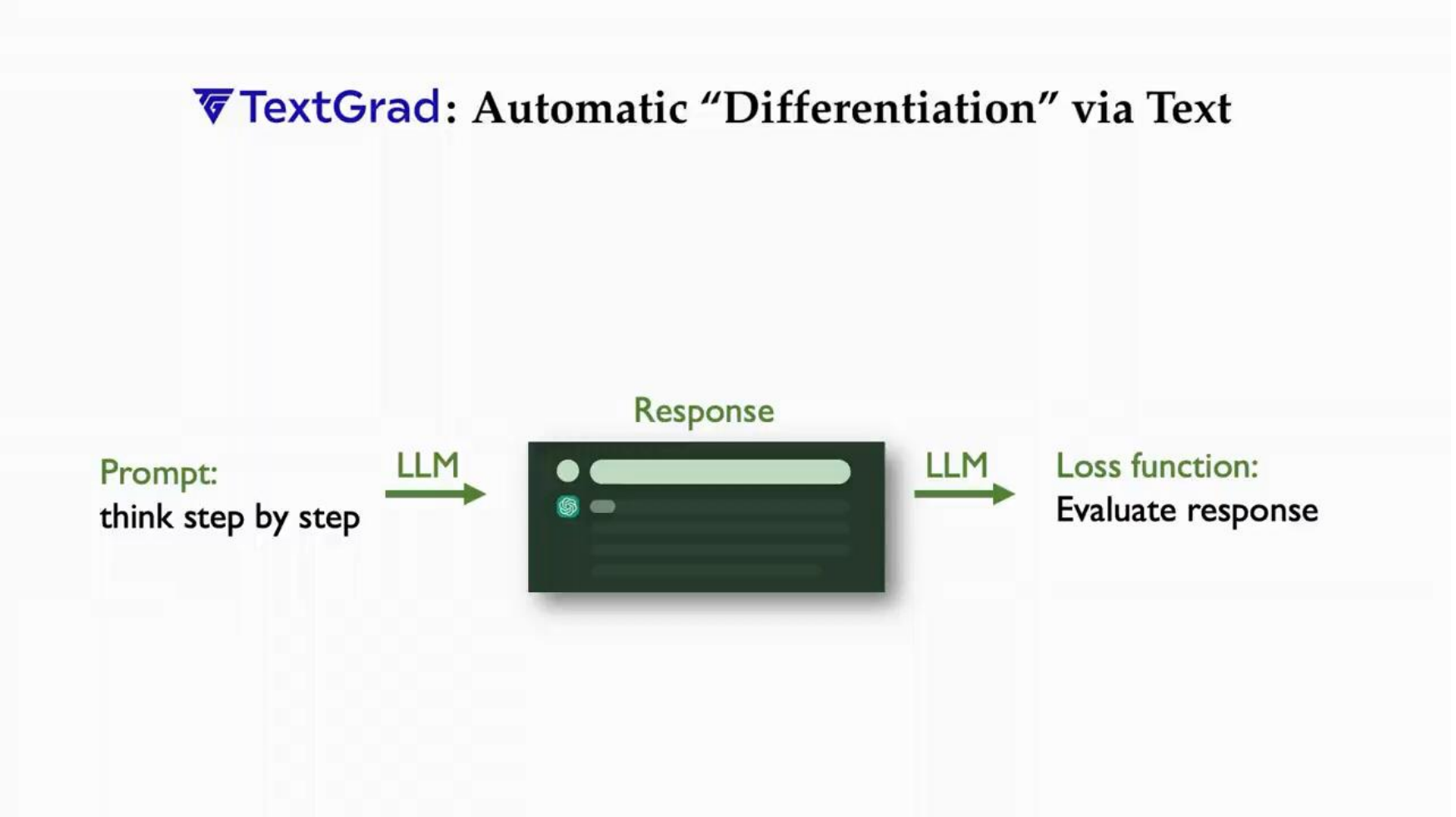
Applications: (i) Non-differentiable loss and inaccessible gradients; (ii) End-to-end optimization: coding, medical report generation, complex reasoning, and molecule design.



FedTextGrad

0. Overview: TextGrad Background

TextGrad Applications: (i) Non-differentiable loss and inaccessible gradients; (ii) End-to-end optimization: coding, medical report generation, complex reasoning, and molecule design.



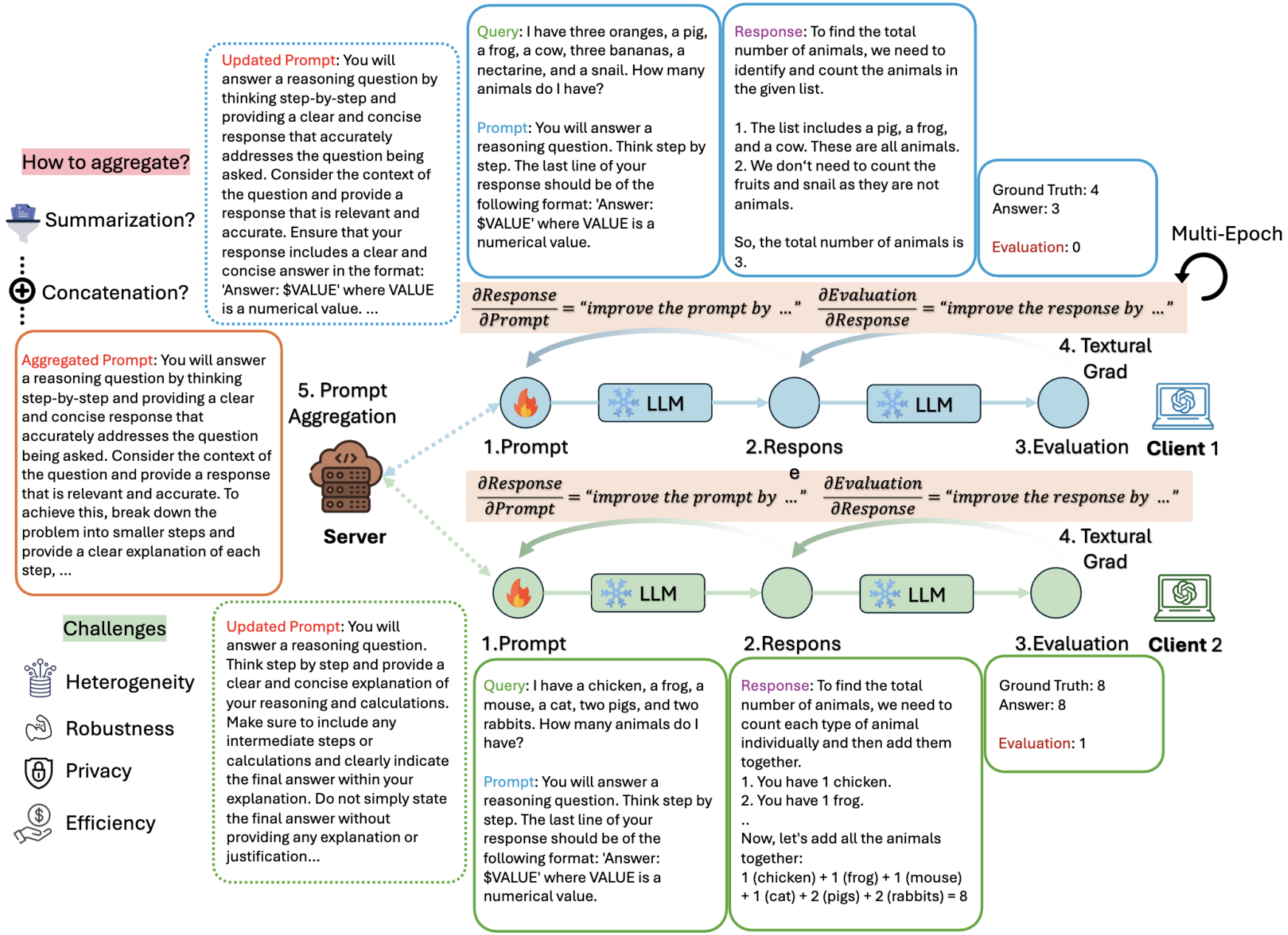
FedTextGrad

0. Overview of FedTextGrad Motivations

A Practical Federated Learning New Setting: Federated TextGrad Framework

Challenges:

Effective and **Efficient** Prompt Aggregation

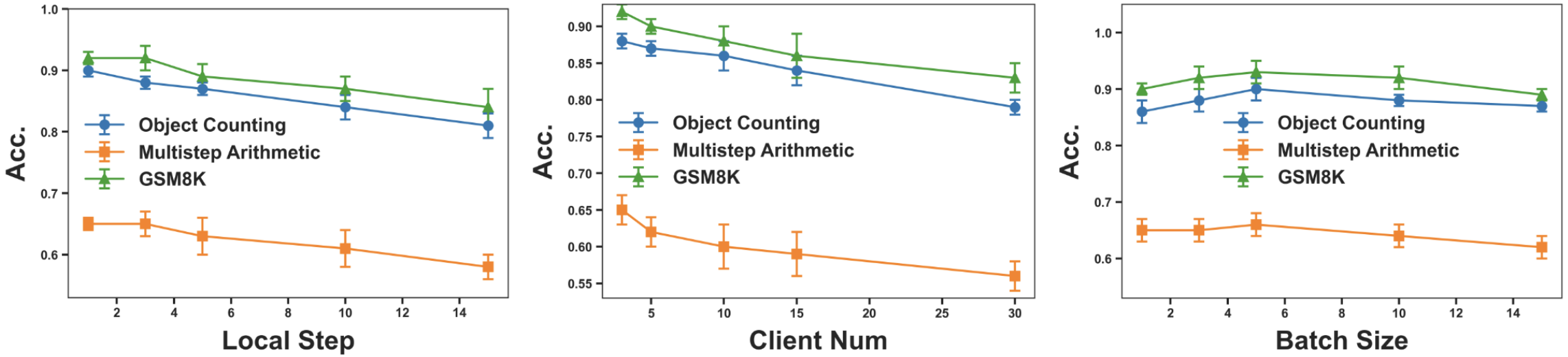


FedTextGrad

3. Experiments

Key Results on Evaluations

Trend Analysis: Increasing the number of local steps and clients results in performance drops, likely due to added heterogeneity and overfitting. Conversely, increasing batch size initially boosts performance but eventually causes a slight decline, indicating an optimal batch size range for maintaining effective prompt optimization.



FedTextGrad

4. Conclusion

Takeaways:

1. **FedTextGrad Introduction**: A novel FL framework leveraging LLM-generated textual gradients for prompt optimization, enabling privacy-preserving collaboration on tasks without numerical loss functions.

2. **Prompt Aggregation Challenges**: Effective aggregation requires balancing length and content; UID-based summarization outperforms concatenation and traditional summarization.

3. **Key Insights**: Optimal local steps and batch sizes enhance performance, while excessive updates or client heterogeneity reduce global alignment. Validated on tasks like BBH and GSM8K.