RIT

Prompting Techniques & Inference Hyperparameters

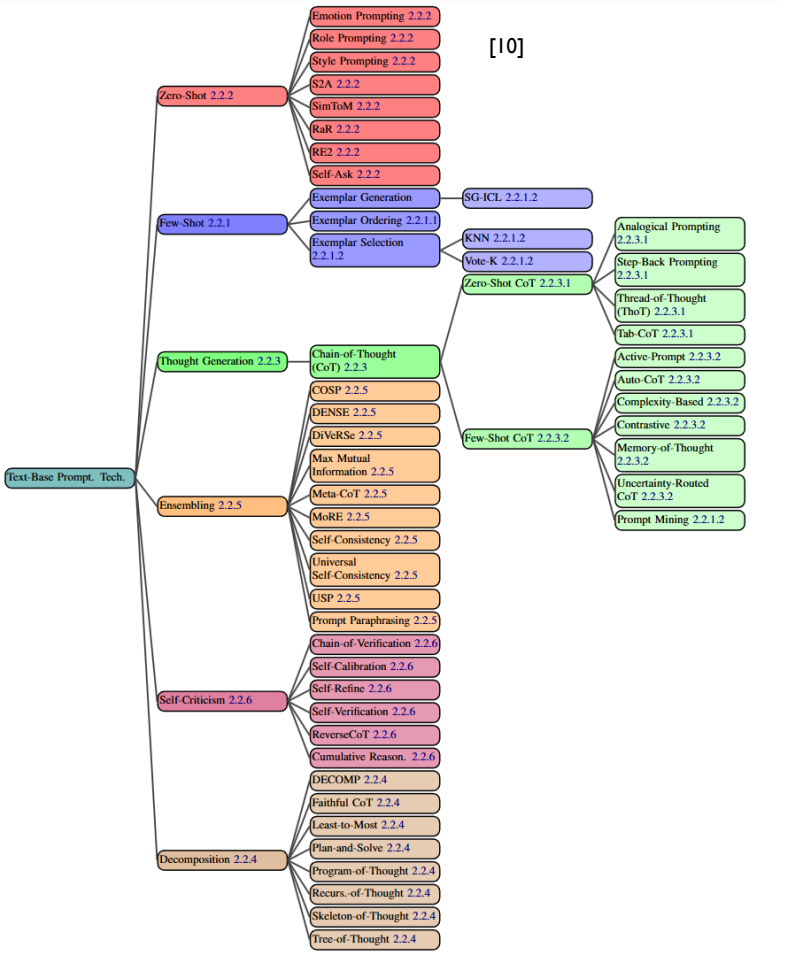
Reza Fayyazi

Prompting Techniques

Prompting Techniques

- Zero-Shot
- Few-Shot
- Chain-of-Thoughts
- Ensembling
- Self-Criticism
- Decomposition

In today's lab, we will work on some of these techniques for cybersecurity operations



[1] Schulhoff, S., Ilie, M., Balepur, N., Kahadze, K., Liu, A., Si, C., ... & Resnik, P. (2024). The Prompt Report: A Systematic Survey of Prompting Techniques. arXiv preprint arXiv:2406.06608.

Zero-Shot

Knowing that <<*MuddyWater has performed credential dumping with Mimikatz and procdump64.exe*>>, what MITRE ATT&CK tactics will a cyber adversary achieve with this technique?

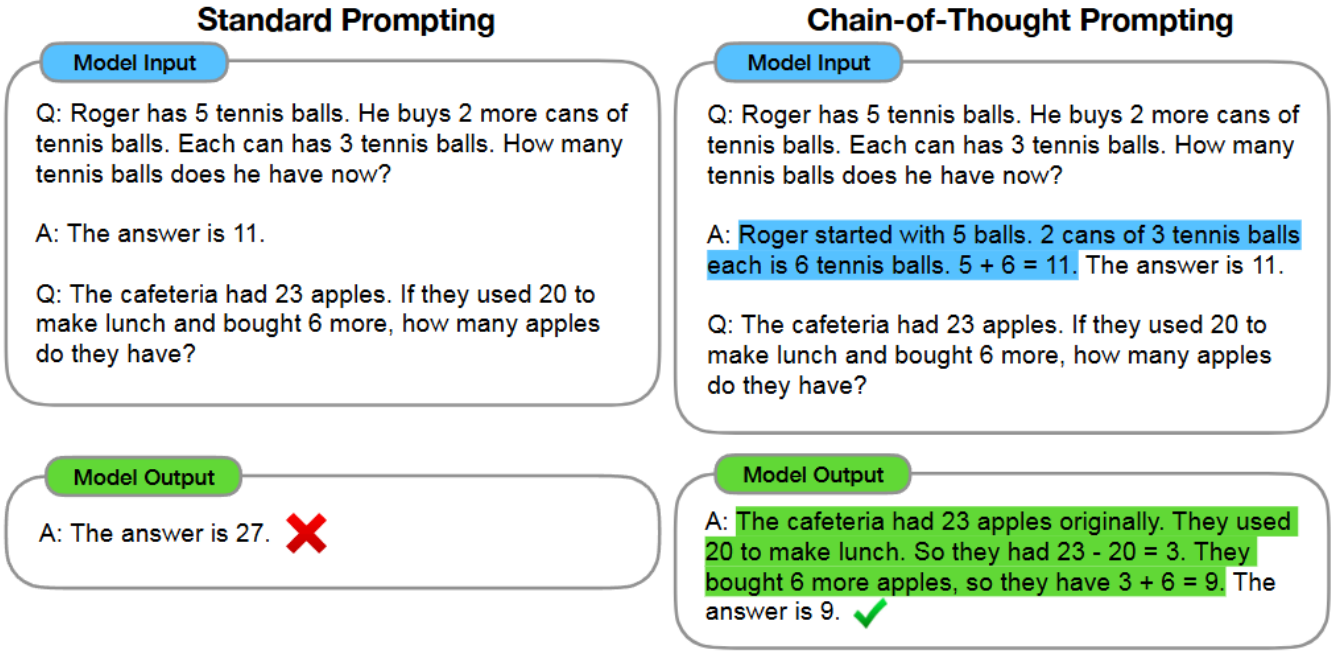
Few-Shot

1. Knowing that <<Adversaries may circumvent mechanisms designed to control elevate privileges to gain higher-level permissions. Most modern systems contain native elevation control mechanisms that …>>, what … ?
Tactic(s): Exfiltration

2. Knowing that <<Adversaries can use stolen session cookies to authenticate to web applications and services>>, what … ?
Tactic(s): Lateral Movement, Defense Evasion

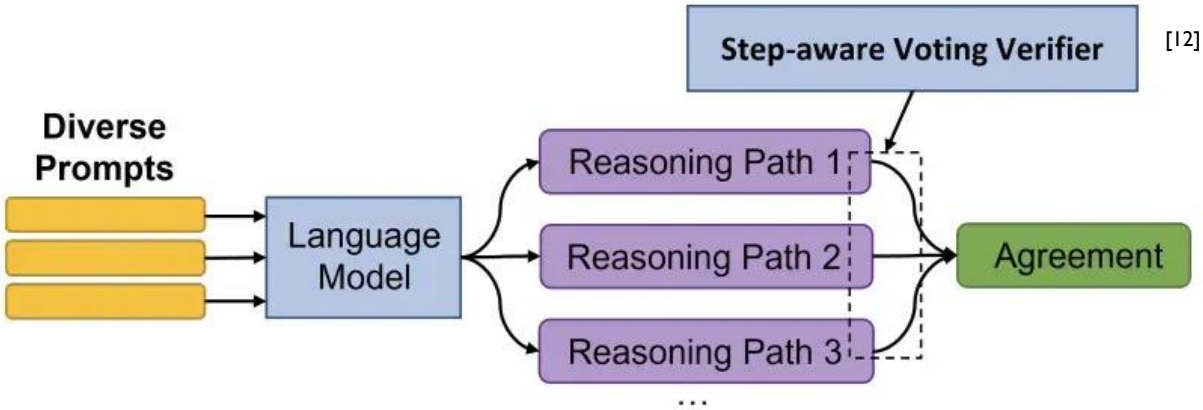
Knowing that <<*MuddyWater has performed credential dumping with Mimikatz and procdump64.exe*>>, what MITRE ATT&CK tactics will a cyber adversary achieve with this technique?

Chain-of-Thoughts [11]



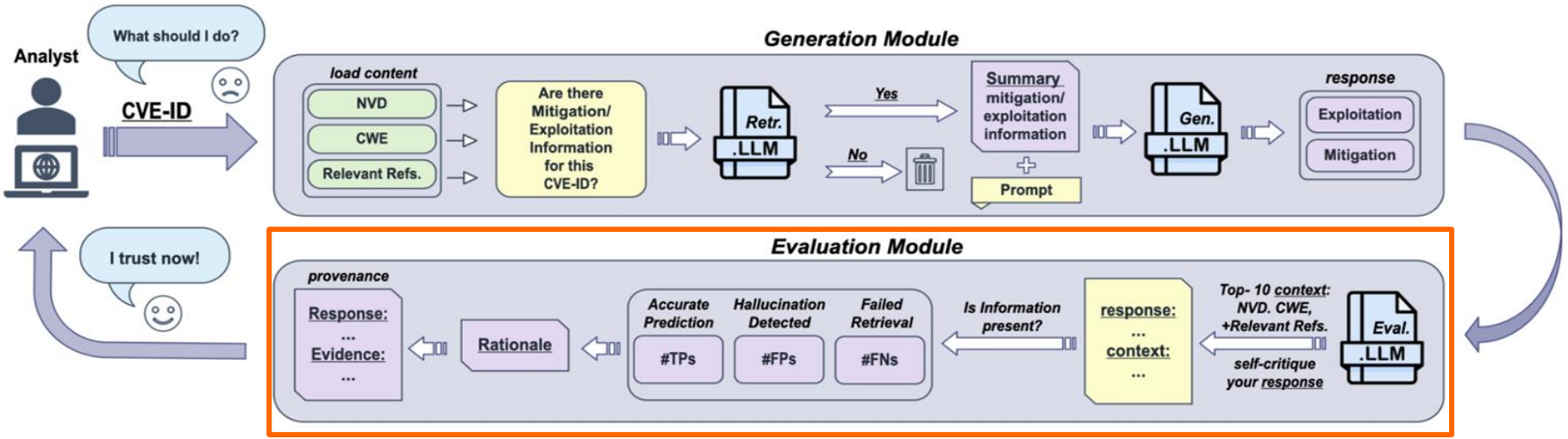
[2] Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., ... & Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. Advances in neural information processing systems, 35, 24824-24837.

Prompt Ensembling



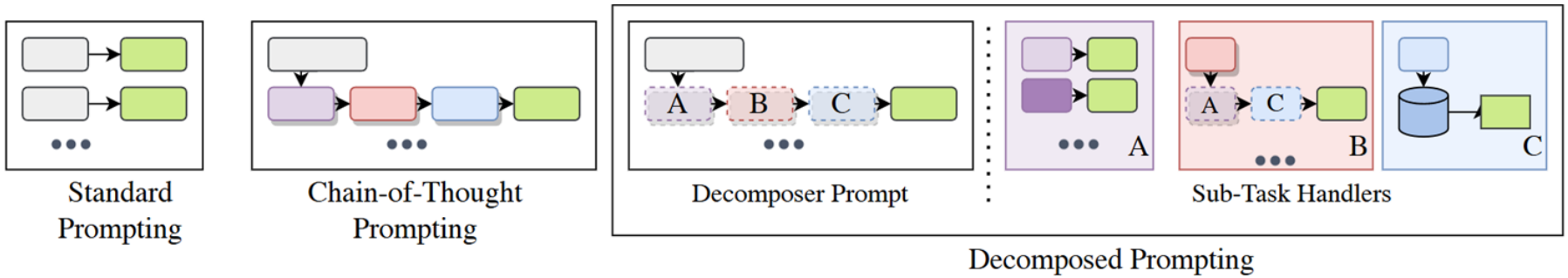
[3] https://learnprompting.org/docs/reliability/ensembling

Self-Criticism – ProveRAG [13]



[4] Fayyazi, R., Trueba, S. H., Zuzak, M., & Yang, S. J. (2024). ProveRAG: Provenance-Driven Vulnerability Analysis with Automated Retrieval-Augmented LLMs. arXiv preprint arXiv:2410.17406.

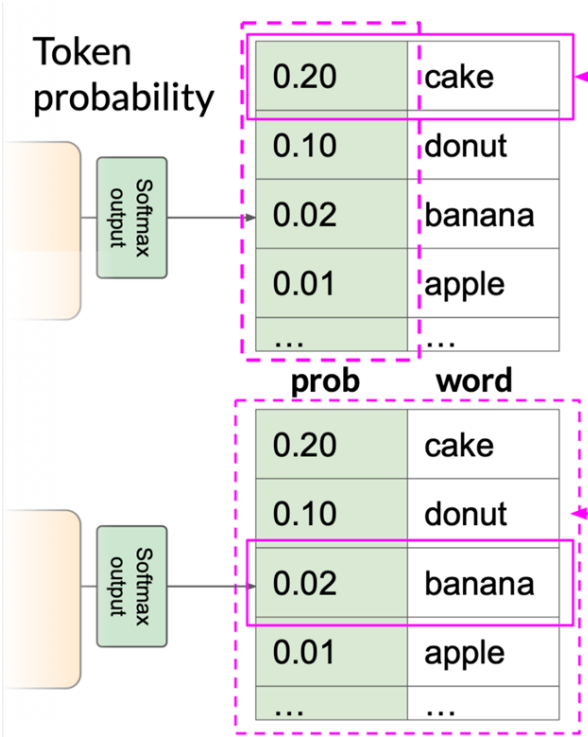
Decomposed Prompting [14]



[5] Khot, T., Trivedi, H., Finlayson, M., Fu, Y., Richardson, K., Clark, P., & Sabharwal, A. Decomposed Prompting: A Modular Approach for Solving Complex Tasks. In The Eleventh International Conference on Learning Representations.

Inference Hyperparameters

Greedy vs. Random Sampling



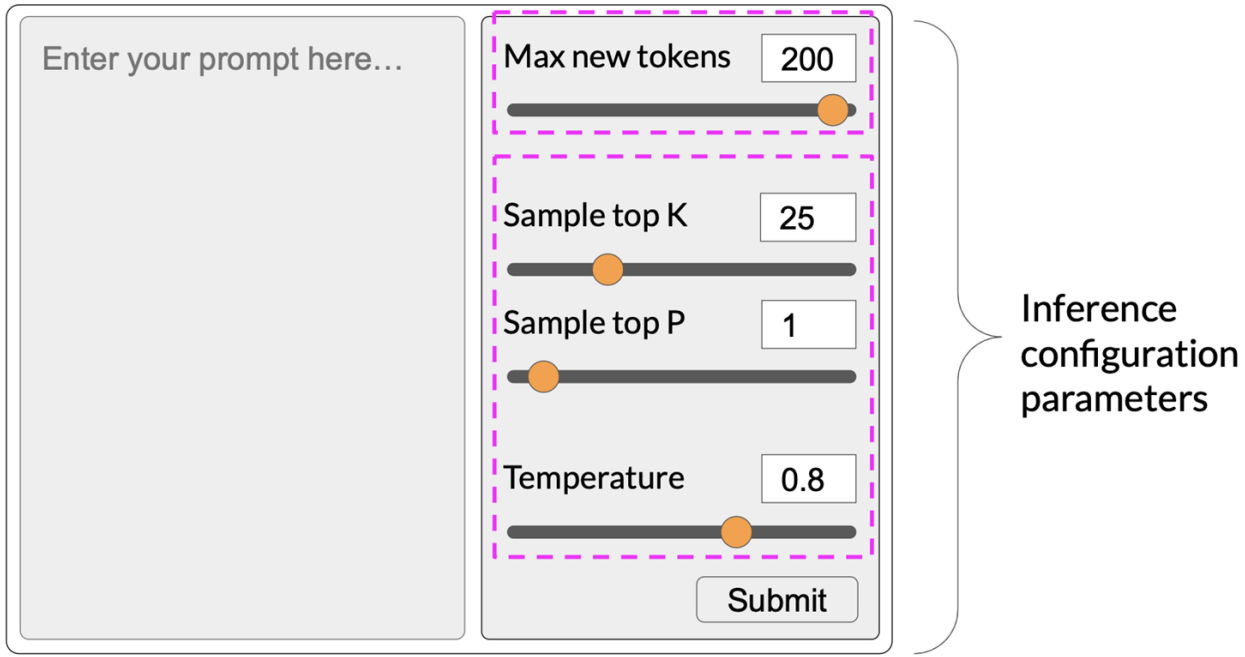
***greedy*: The word/token with the highest probability is selected.**

***random(-weighted) sampling*: select a token using a random-weighted strategy across the probabilities of all tokens.**

Here, there is a 20% chance that 'cake' will be selected, but 'banana' was actually selected.

[6] https://www.coursera.org/learn/generative-ai-with-llms/

Inference Hyperparameters



[6] https://www.coursera.org/learn/generative-ai-with-llms/

Thank you!

Reza Fayyazi
rf1679@rit.edu