

Shubhra Kanti Karmaker (“Santu”), Dongji Feng
Auburn University, BDI Lab

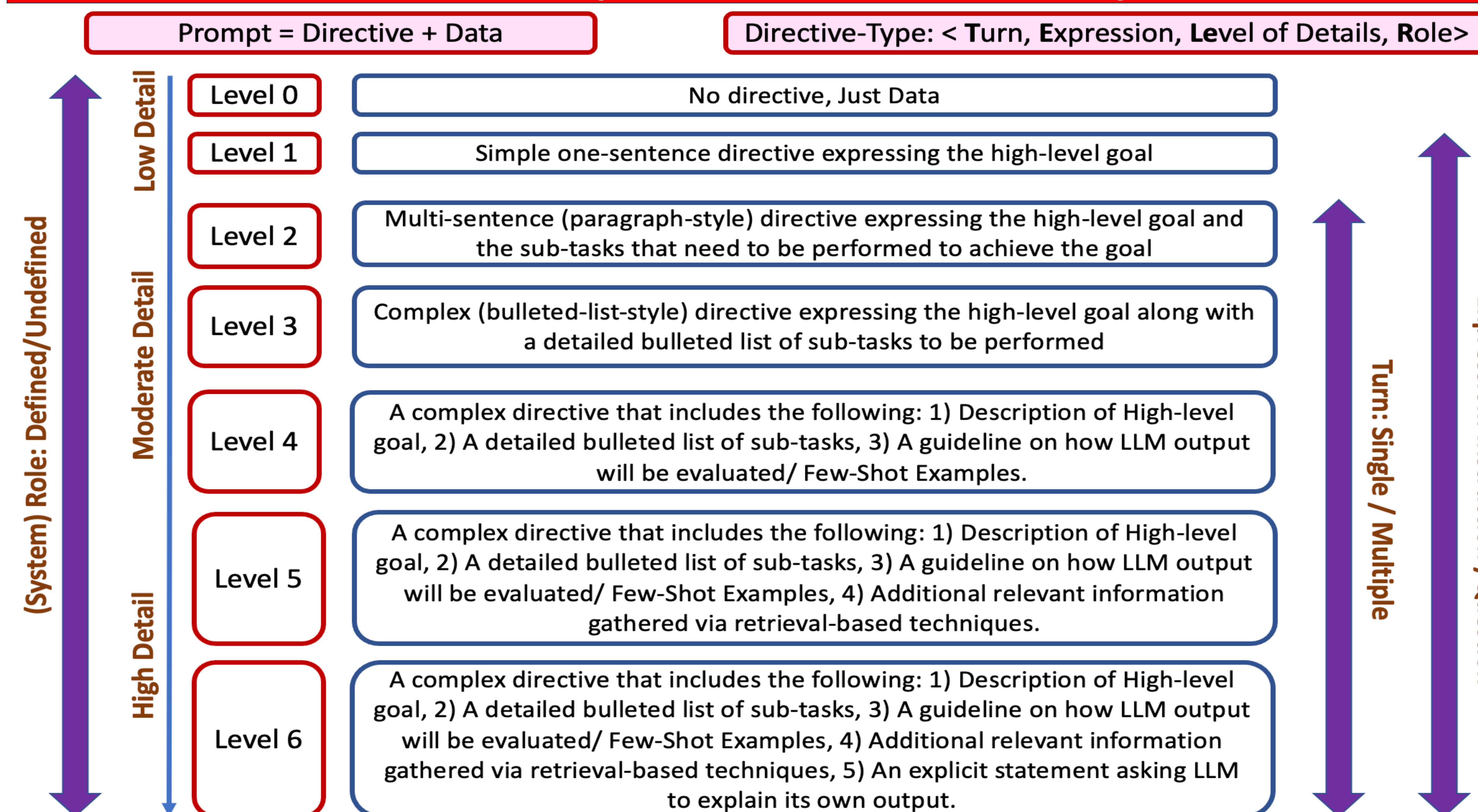
Summary

- We proposed **TELeR** (Turn, Expression, Level of Details, Role)
- A general Taxonomy can be used to design prompts for **complex task**.
 - Enabling **meaningful comparisons** across different prompt designs.
 - Draw more **accurate** conclusion about LLM’s performance.

Motivation

- **Quality** of prompts greatly influence the performance of LLMs.
- Studies with diverse prompts for ill-defined **complex tasks** are still rare.
- Different prompt **types/styles** are used and different **degrees of details** provided in prompts.
- **Apple-to-apple comparison** not made in prompts comparison
- **Goal:** Establishing a common standard through this taxonomy for accurate conclusion about LLM’s performance.

Proposed Taxonomy



Use case Study

Meta-Review Generation

- Meta reviewing is a **complex task** involves summarizing expert reviews from multiple reviewers.
- We assume three comments (R_1, R_2, R_3) as *data* for this task.
- We only provide prompt with **Level 3, 4** due to the space.

Level 3 Prompt: *Prepare a meta-review by answering the following questions from the reviewer comments (provided after the questions).*

1. *Based on the reviewer’s comments, what are the core contributions made by the authors?*
2. *What are the common strengths of this work, as mentioned by multiple reviewers?*
3. *What are the common weaknesses of this work, as highlighted by multiple reviewers?*
4. *What suggestions would you provide for improving this paper?*
5. *What are the missing references mentioned by the individual reviews?*

The review texts are below: $\langle R_1, R_2, R_3 \rangle$

Level 4 Prompt: “Level 3 Prompt” + “A good output should be coherent, highlight major strengths/issues mentioned by multiple reviewers, be less than 400 words in length, and finally, the response should be in English only”.

