"...There is a gentle river flowing down the forest. On the left side of the river stands a bear wearing a shirt and a hat, while on the right is a mushroom..."

Text-to-Image

Stable Diffusion   Attend-and-excite

DALL·E 2   Structured Diffusion

Object Missing & Wrong Relation

LLM

...a gentle *river* flows **down** the forest...

...**bear wearing a shirt and a hat** on the **left**...

...a **mushroom** on the **right**...

Step 1   Step 2   Step 3

User Inference

✏ "change shirt to **blue shirt**"   🗑 "Delete the **hat**"

Our Progressive Text-to-Image