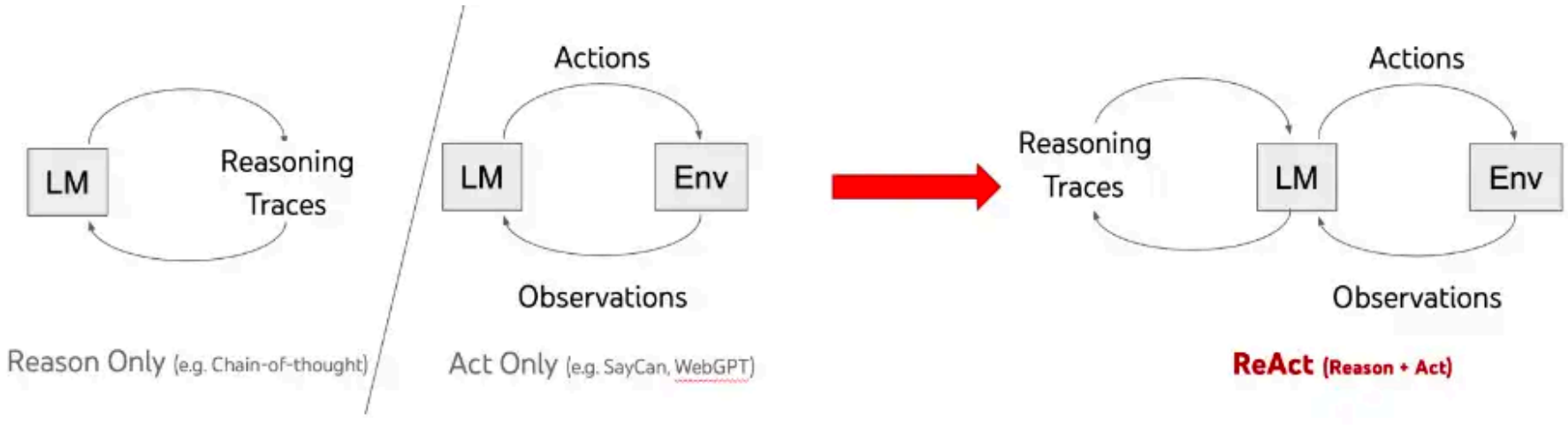
Analytics Vidhya

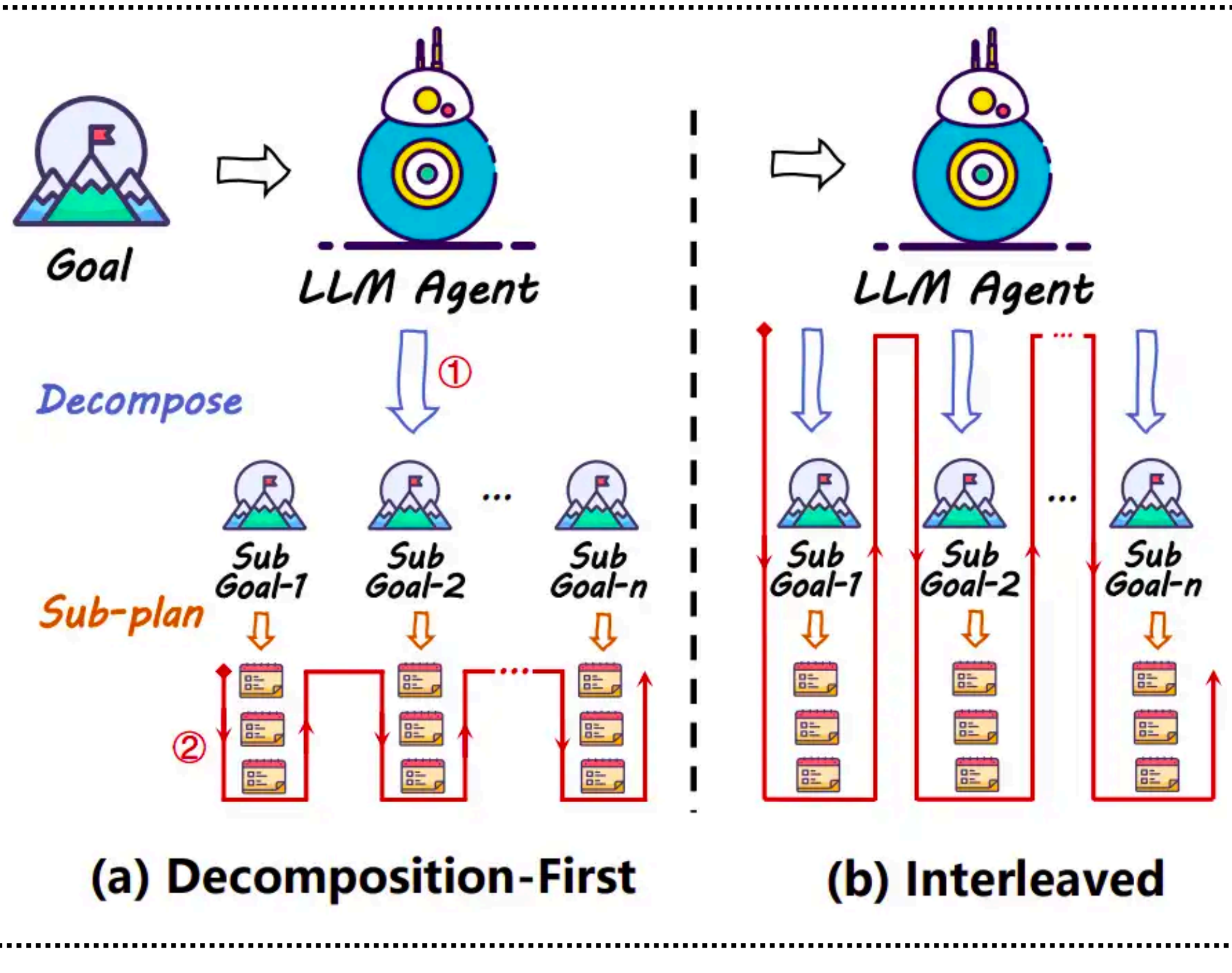
Day 8 of Mastering AI Agents

Agentic AI Planning Pattern



Language models are getting better at reasoning (e.g. chain-of-thought prompting) and acting (e.g. WebGPT, SayCan, ACT-1), but these two directions have remained separate.

ReAct asks, what if these two fundamental capabilities are combined?



Agentic AI Planning Pattern

The Agentic AI Planning Pattern is a framework that focuses on breaking down a larger problem into smaller tasks, managing those tasks effectively, and ensuring continuous improvement or adaptation based on task outcomes. The process is iterative and relies on a structured flow to ensure that the AI system can adjust its plan as needed, moving closer to the desired goal with each iteration.
The Planning Pattern has the following main components:

Planning

- In this initial stage, the AI agent interprets the prompt and devises an overall plan.
- The plan outlines how the AI intends to tackle the problem, including high-level goals and strategies.

Generate Task

- From the plan, the AI system generates specific tasks that must be executed.
- Each task represents a smaller, manageable portion of the overarching goal, allowing the AI to work in focused steps.

Single Task Agent

- The Single Task Agent is responsible for completing each task generated in the previous step.
- This agent executes each task using predefined methods like ReAct (Reason + Act) or ReWOo (Reasoning WithOut Observation).
- Once a task is completed, the agent returns a Task Result, which is sent back to the planning loop.

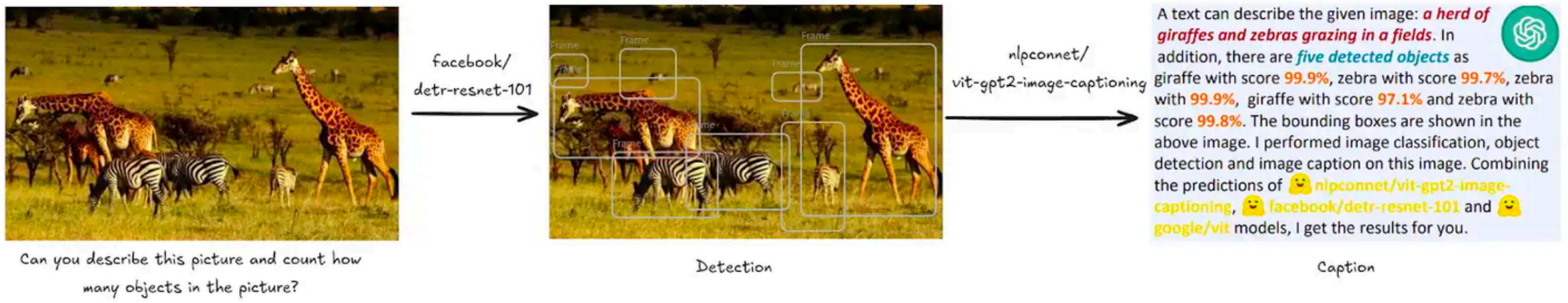
Replan

- The Replan stage evaluates the Task Result to determine if any adjustments are needed.
- If the task execution does not fully meet the desired outcome, the system will replan and possibly modify the tasks or strategies.
- This feedback loop allows the AI system to learn and improve its approach iteratively, making it more adaptable to changing requirements or unexpected outcomes.

Iterate

- This part of the pattern is a loop connecting Generate Task and Replan.
- It signifies the iterative nature of the process, where the AI system continuously re-evaluates and adjusts its approach until it achieves satisfactory results.

Example of an Agentic AI Planning Pattern



The above-given illustration depicts a sequential image understanding process, with steps that align with the agentic AI planning pattern. In agentic AI, an “agent” takes actions based on observations and planned responses to achieve a specific goal. Here’s how each step in the image fits into the agentic AI framework:

Goal Setting (Understanding the Task)

- **Prompt**: The task begins with a question: “Can you describe this picture and count how many objects are in the picture?”

- **Agentic AI Element**: The AI agent interprets this goal as a directive to analyze the image for both object recognition and description. The goal is to answer the question comprehensively by identifying, counting, and describing objects.

Planning and Subgoal Formation

Process Breakdown:

To accomplish this goal, the agent breaks the task down into specific subtasks:

- Object Detection (identify and localize objects)
- Classification (identify what each object is)
- Caption Generation (generate a natural language description of the scene)

Agentic AI Element: An agent plans its actions by setting intermediate subgoals in the agentic AI planning pattern. Here, detecting objects is a subgoal required to complete the ultimate objective (generating a descriptive caption that includes a count of objects).

Perception and Action (Detecting and Describing)

Tools and Models Used

- The agent utilises the facebook/detr-resnet-101 model for detection, which identifies and locates objects (e.g., giraffes and zebras) and assigns confidence scores.
- After detection, the agent uses nlpconnect/vit-gpt2-image-captioning to generate a descriptive caption.

- **Agentic AI Element**: The agent “perceives” its environment (the image) using specific perception modules (pre-trained models) that allow it to gather necessary information. In agentic AI, perception is an active, goal-oriented process. Here, the models act as perception tools, processing visual information to achieve the overall objective.

Evaluation and Iteration (Combining Results)

- **Processing and Aggregating Information**: The results from detection (bounding boxes and object types) and captioning (descriptive text) are combined. The agent evaluates its outputs, confirming both object detection confidence levels and the coherence of the description.

- **Agentic AI Element**: Agentic AI involves continuously evaluating and adjusting responses based on feedback and information aggregation. The agent reviews its predictions (detection scores and bounding boxes) to ensure they align with the task’s demands.

Goal Achievement (Answer Presentation)

- **Output Presentation**: The agent finally provides an answer that includes a count of detected objects, a list of identified objects with confidence scores, and a descriptive caption.

- **Agentic AI Element**: The agent completes the goal by synthesising its perception and planning outcomes into a coherent response. In agentic AI, this step is about achieving the task's overarching goal and generating an output that addresses the user's initial question.

For more information, kindly visit this article