

ReAct agent

Graph Prototype and Agent Evolution

Aleš Kalfas

June 27, 2024 v3



ReAct agent *outlines*

1. Graph prototype
 1. Motivation
 2. Example
 3. Formalization + prototype
2. Evolution
 1. Critic
 2. Replan
 3. Human in the loop
 4. Human in the action
 5. Replan example
 6. (TBD) Reflexion



ReAct agent *graph prototype*

Motivation

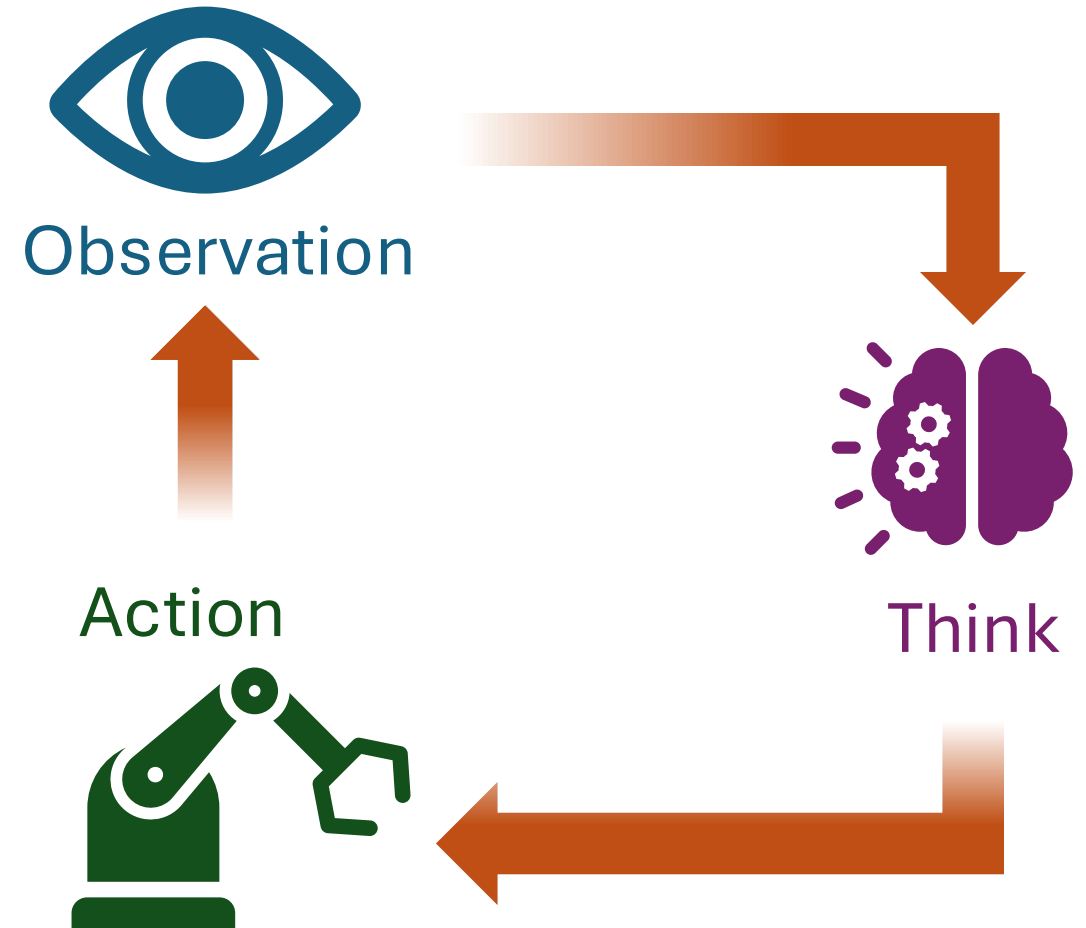


ReAct agent *graph prototype*

Motivation

Why ReAct agent?

- ReAct agent is a basically the minimal implementation of the general agentic behavior
- LLM is used to generate reasoning traces and task-specific actions (Thinking)
- The action step allows to interact with external world (Observation / Action)
- LLM controlled inner loop





ReAct agent *graph prototype*

Minimal graph



ReAct agent *graph prototype*

Minimal graph

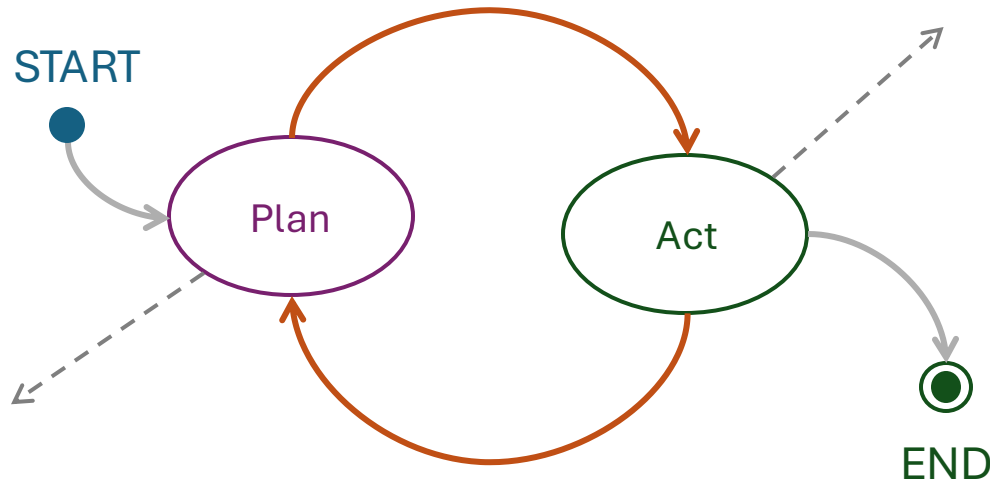


Take an action

- Calculator
- Search-place
- Weather-forecast
- Finish



Call LLM



```
export enum Status {
  UNCOMPLETED = "uncompleted",
  COMPLETED = "completed",
}

export interface ActionDefinition {
  name: string;
  description: string;
  args: any;
}

export interface ActionCall {
  name: string;
  args: any;
}

export interface Step {
  thought: string;
  action: ActionCall;
  action_output?: string;
}

export interface Message {
  role: "user" | "assistant";
  content: string;
  created_at: Date;
}

export interface State {
  messages: Message[];
  input: string;
  question?: string;
  output?: string;
  status: Status;
  steps: Step[];
}
```



ReAct agent *graph prototype*

Minimal graph + Critic

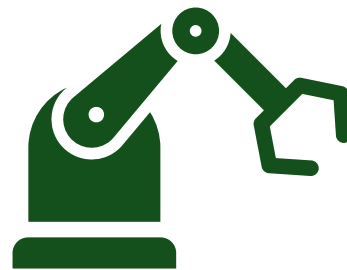


ReAct agent *graph prototype*

Minimal graph + Critic



Call LLM



Take an action

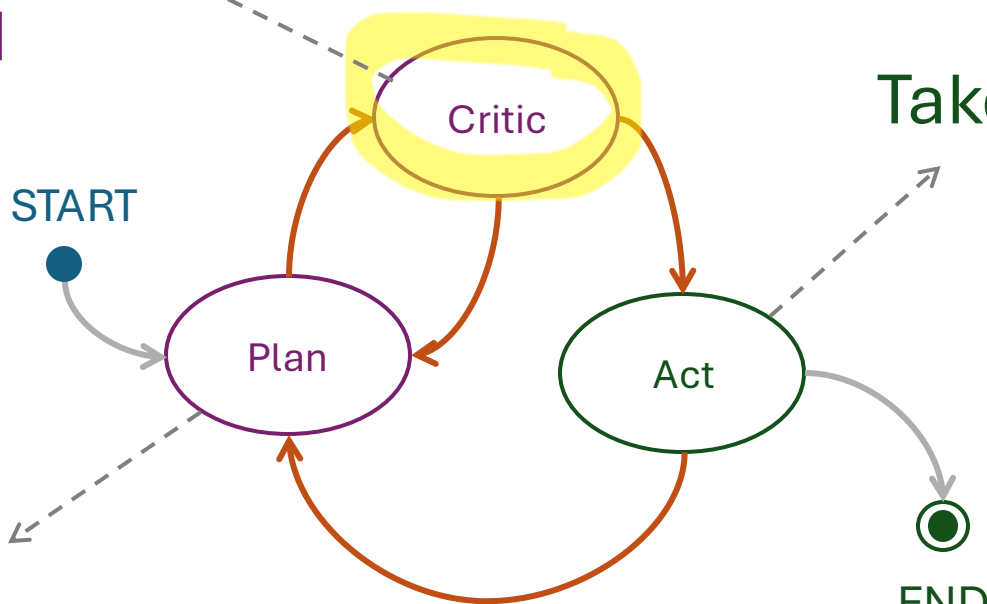
- Calculator
- Search-place
- Weather-forecast
- Finish

START

END



Call LLM



```
export enum Status {  
  UNCOMPLETED = "uncompleted",  
  COMPLETED = "completed",  
}
```

```
export interface ActionDefinition {  
  name: string;  
  description: string;  
  args: any;  
}
```

```
export interface ActionCall {  
  name: string;  
  args: any;  
}
```

```
export interface Step {  
  thought: string;  
  action: ActionCall;  
  action_output?: string;  
  criticism?: string;  
  excluded: boolean;  
}
```

```
export interface Message {  
  role: "user" | "assistant";  
  content: string;  
  created_at: Date;  
}
```

```
export interface State {  
  messages: Message[];  
  input: string;  
  question?: string;  
  output?: string;  
  status: Status;  
  steps: Step[];  
}
```



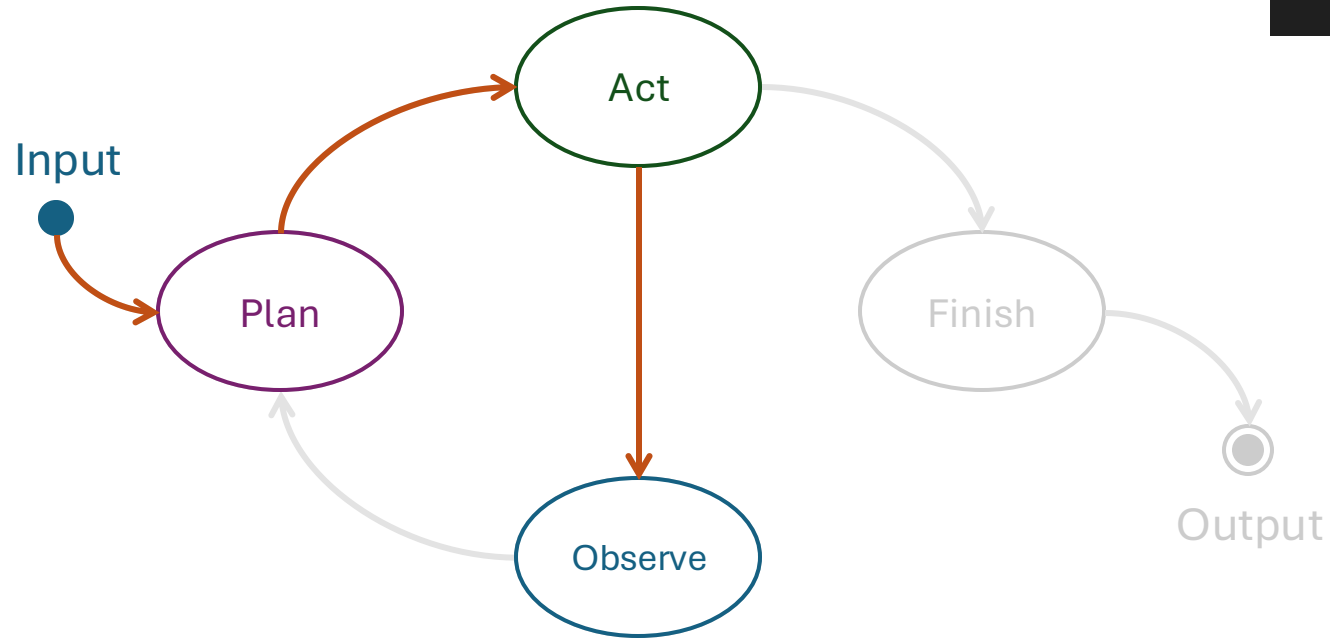

ReAct agent *graph prototype*

Example



ReAct agent *graph prototype*

Example



Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

Step[1]

thought: "I need to know the area of the UK."
action: "Wikipedia[Area of the UK]"
action_output: "The United Kingdom of Great Britain and Northern Ireland..."
observation: "UK area is 244,376 km²"

Chat



How many times does the UK fit into the USA?

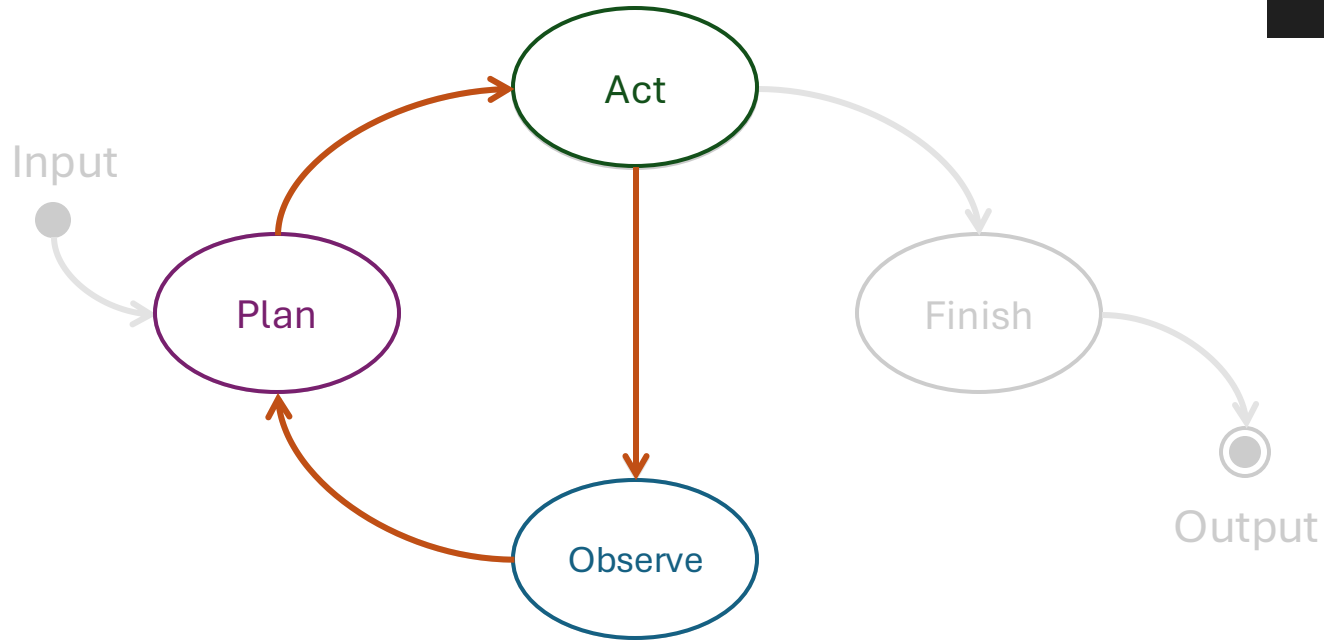


Thinking...



ReAct agent *graph prototype*

Example



Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

Step[1]

thought: "I need to know the area of the UK."
action: "Wikipedia[Area of the UK]"
action_output: "The United Kingdom of Great Britain and Northern Ireland..."
observation: "UK area is 244,376 km²"

Step[2]

thought: "I need to know the area of the US."
action: "Wikipedia[Area of the US]"
action_output: "The United States of America (USA or U.S.A.), commonly known as..."
observation: "US area is approximately 9,833,520 km²"

Chat



How many times does the UK fit into the USA?

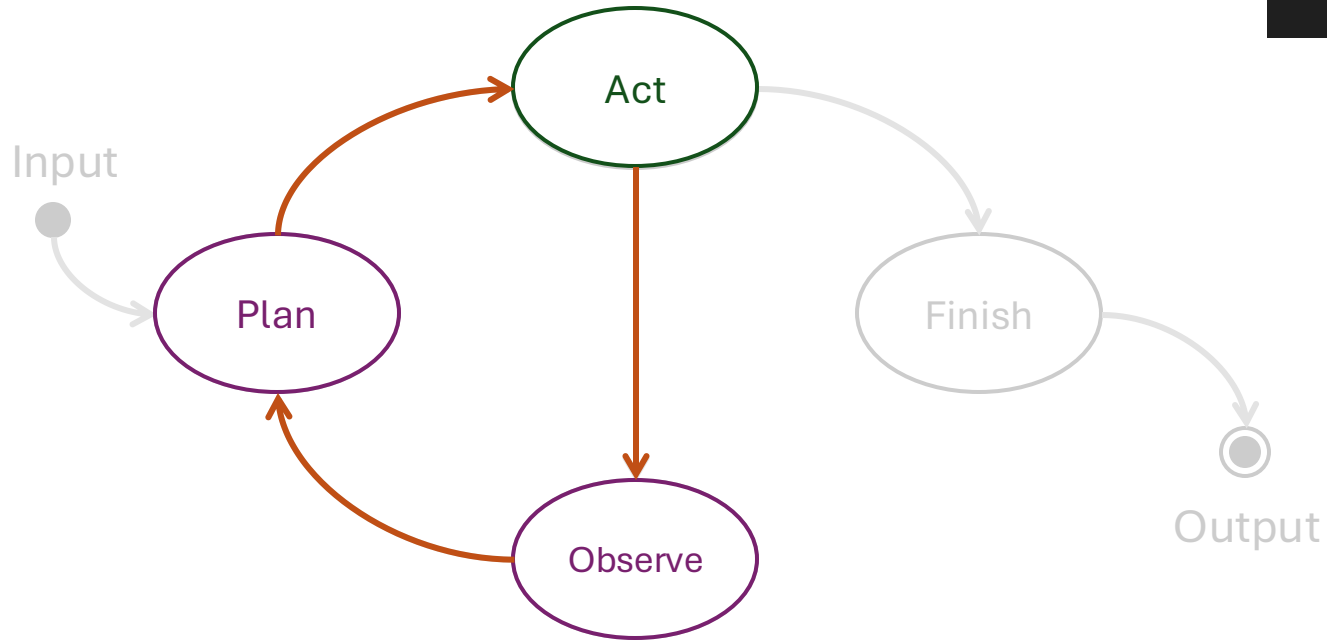


Thinking...



ReAct agent *graph prototype*

Example



Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

Step[1]

thought: "I need to know the area of the UK."
action: "Wikipedia[Area of the UK]"
action_output: "The United Kingdom of Great Britain and Northern Ireland..."
observation: "UK area is 244,376 km2"

Step[2]

thought: "I need to know the area of the US."
action: "Wikipedia[Area of the US]"
action_output: "The United States of America (USA or U.S.A.), commonly known as..."
observation: "US area is approximately 9,833,520 km2"

Step[3]

thought: "I can calculate how many times the UK fits into the USA."
action: "Python[Calculating the ratio of the areas]"
action_output: "40.2393033685"
observation: "The result of the ratio is 40.2393033685."

Chat



How many times does the UK fit into the USA?

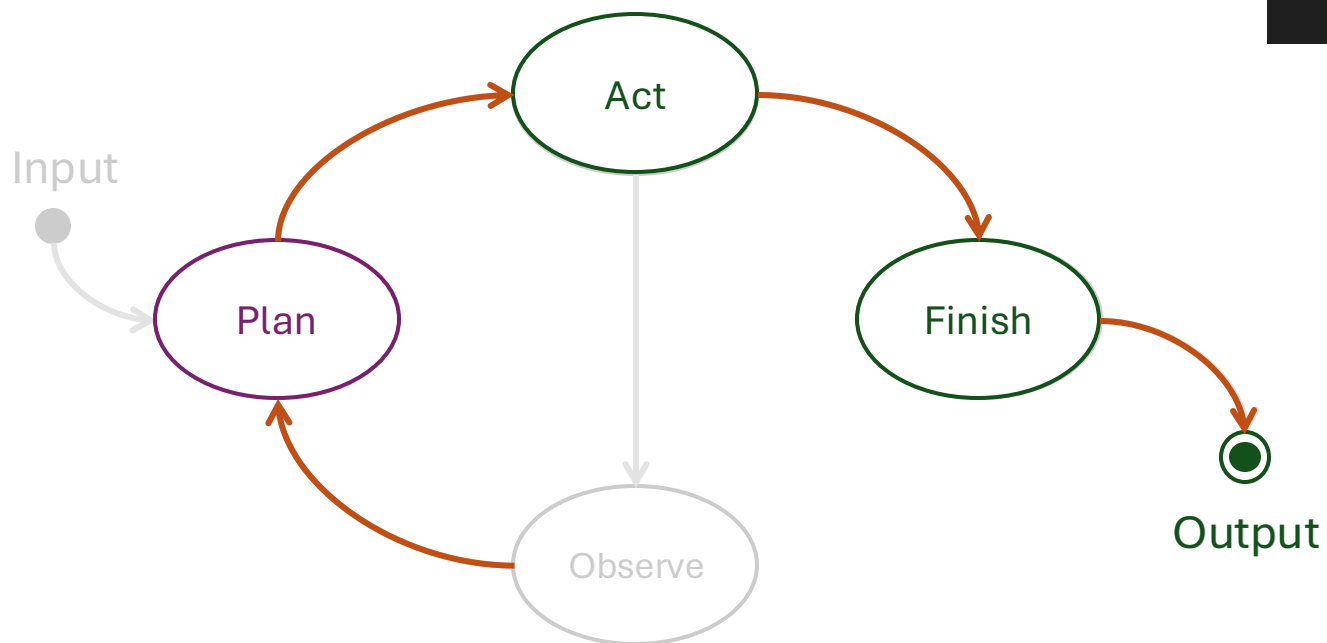


Thinking...



ReAct agent *graph prototype*

Example



Task

question: "How many times does the UK fit into the USA?"

state: Completed

answer: "Final[Approximately 40.23 times]"

Step[1]

thought: "I need to know the area of the UK."

action: "Wikipedia[Area of the UK]"

action_output: "The United Kingdom of Great Britain and Northern Ireland..."

observation: "UK area is 244,376 km²"

Step[2]

thought: "I need to know the area of the US."

action: "Wikipedia[Area of the US]"

action_output: "The United States of America (USA or U.S.A.), commonly known as..."

observation: "US area is approximately 9,833,520 km²"

Step[3]

thought: "I can calculate how many times the UK fits into the USA."

action: "Python[Calculating the ratio of the areas]"

action_output: "40.2393033685"

observation: "The result of the ratio is 40.2393033685."

Step[4]

thought: "I have the final answer."

action: "Final[Approximately 40.23 times]"

Chat



How many times does the UK fit into the USA?



Thinking...



Approximately 40.23 times



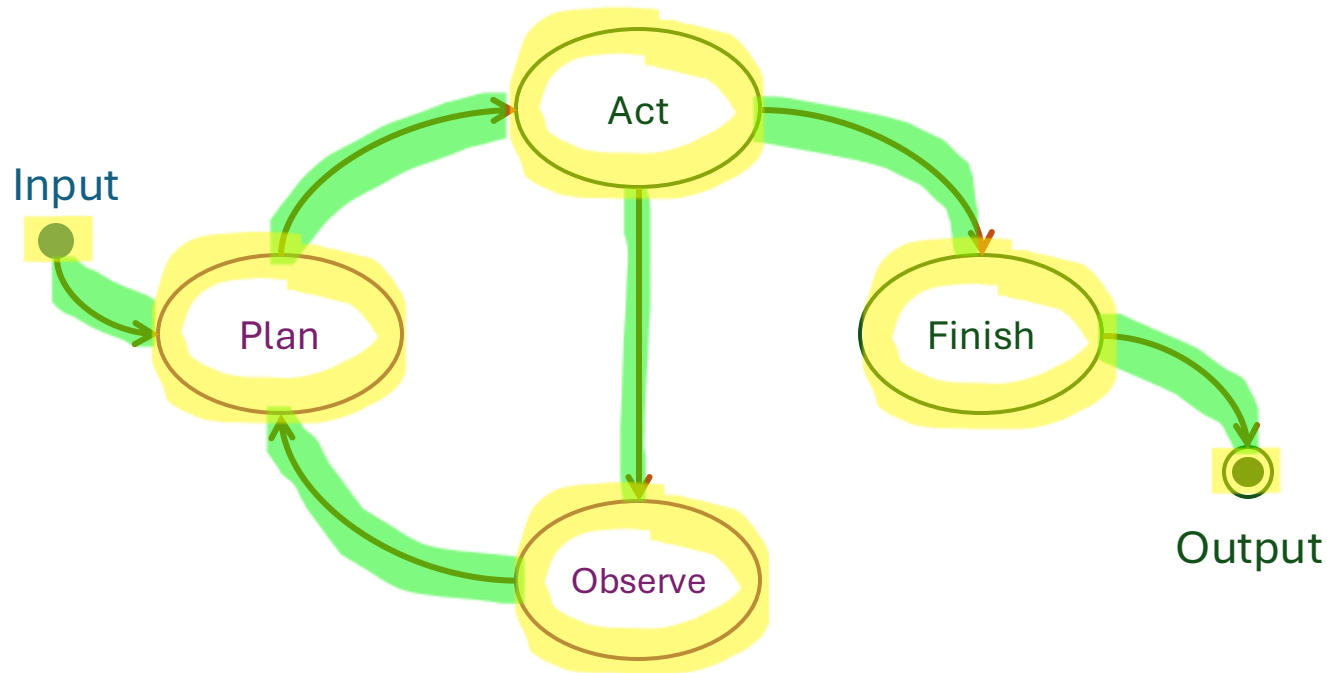
ReAct agent *graph prototype*

Formalization



ReAct agent *graph prototype*

Formalization



Graph

- Structure

- **Nodes**

- Nodes are executable units interconnected by edges.
 - Each graph should contain at least two nodes **Input** and **Output**

- **Edges**

- Edges determine the order of executed nodes.

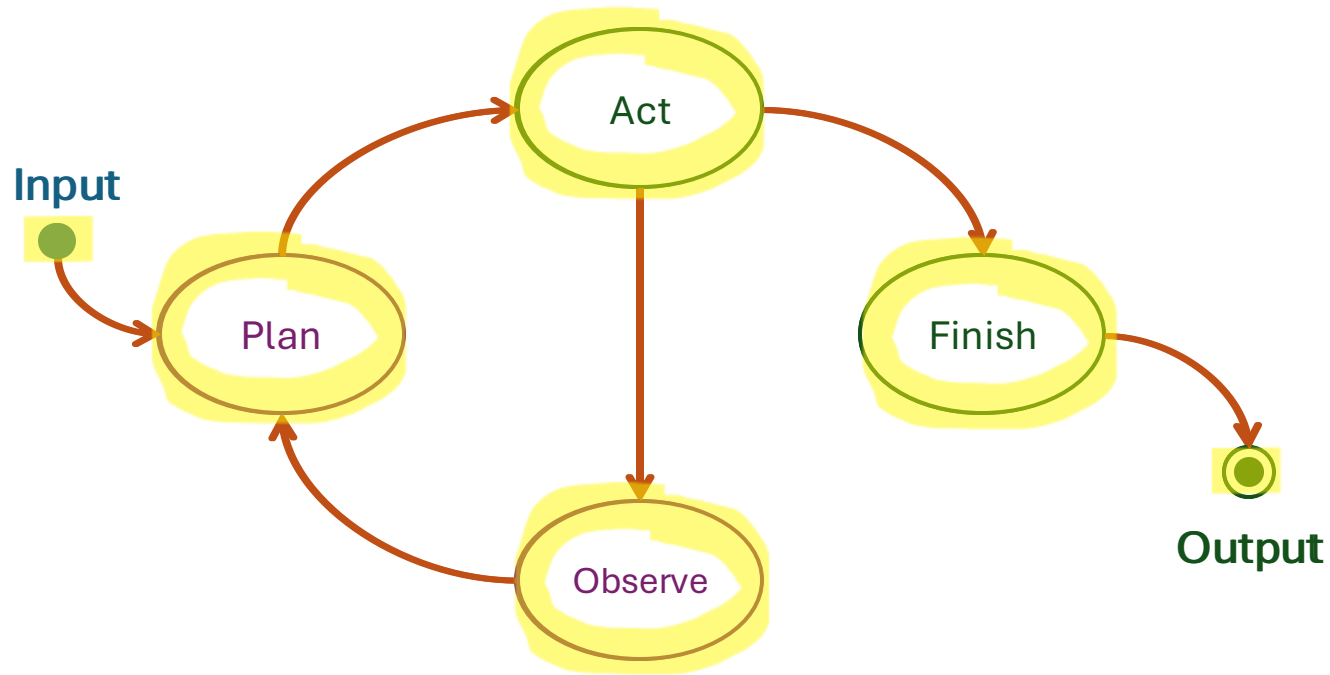
- Behavior

- Graph execution contains do-while loop which traverse through nodes due to edges and their orientation.
 - Execution starts in the **Input** node and ends in the **Output** node.



ReAct agent *graph prototype*

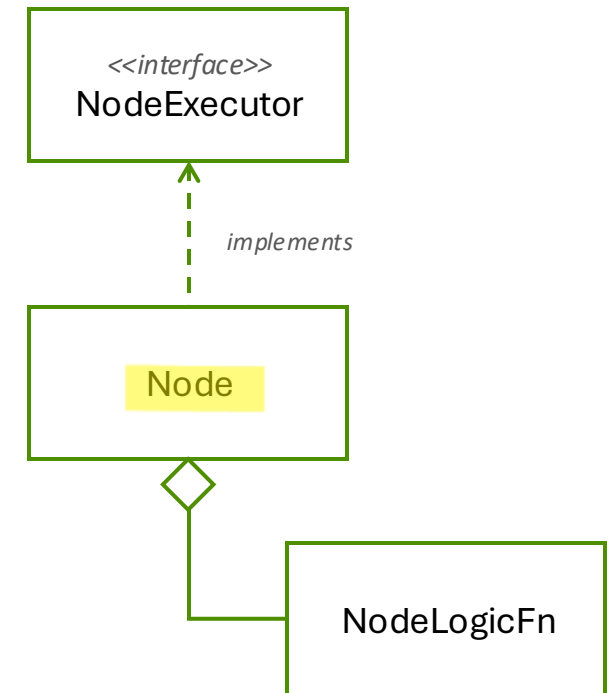
Formalization



- *class Node*

- Each **Node** is an executable unit that implements **NodeExecutor**. It has just a one method the **execute**.
- It receives a **logic:NodeLogicFn** function on its creation. This function is processed when **Node.execute** is called.

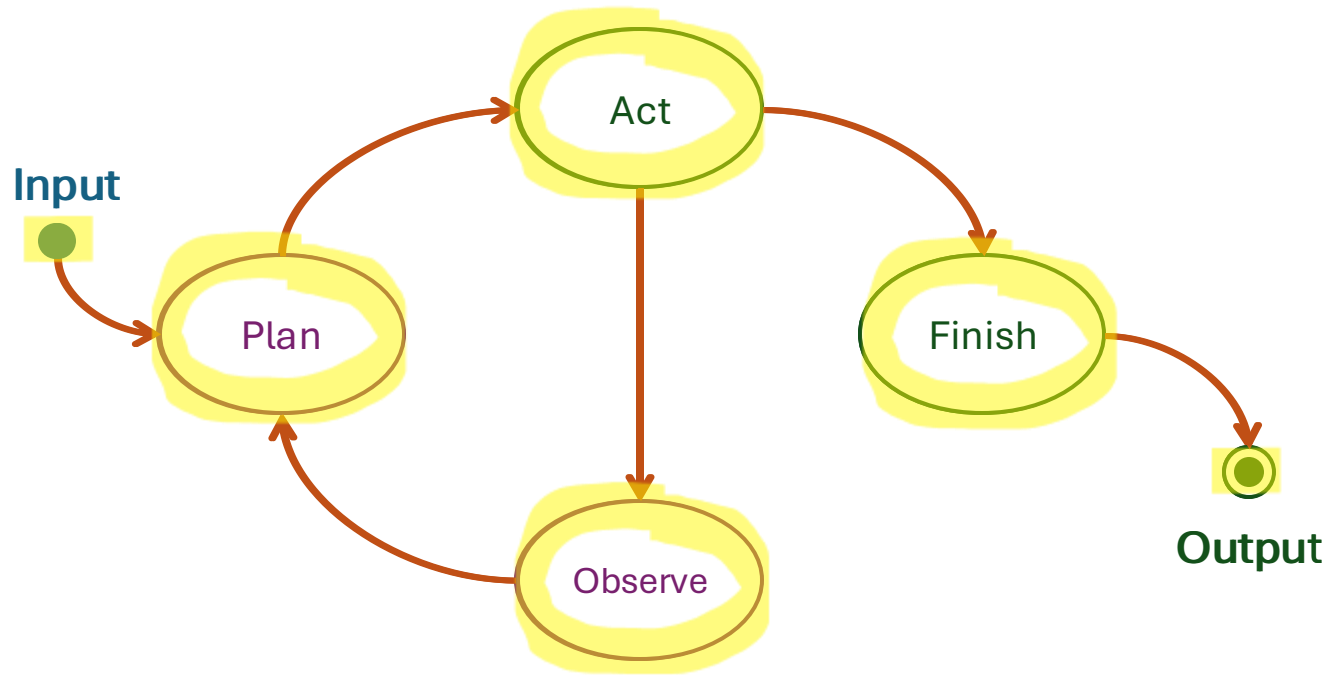
Class diagram





ReAct agent *graph prototype*

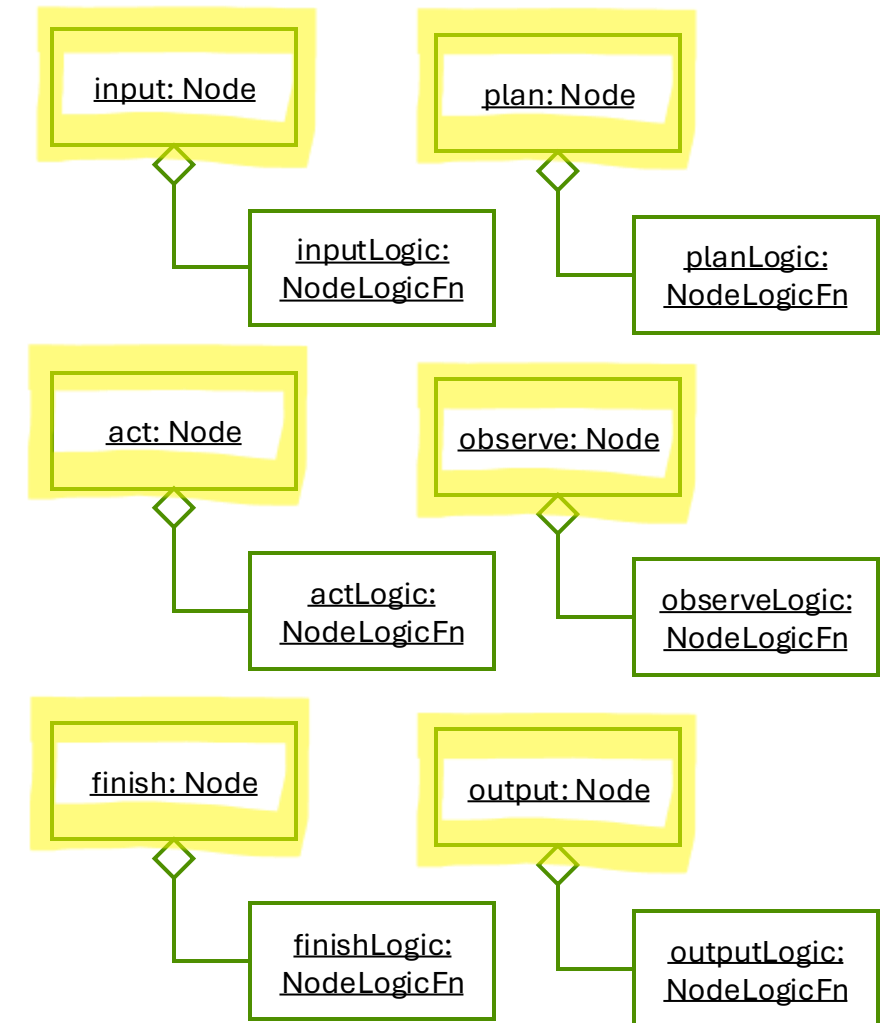
Formalization



- *class Node*

- Each **Node** is an executable unit that implements **NodeExecutor**. It has just a one method the **execute**.
- It receives a **logic:NodeLogicFn** function on its creation. This function is processed when **Node.execute** is called.

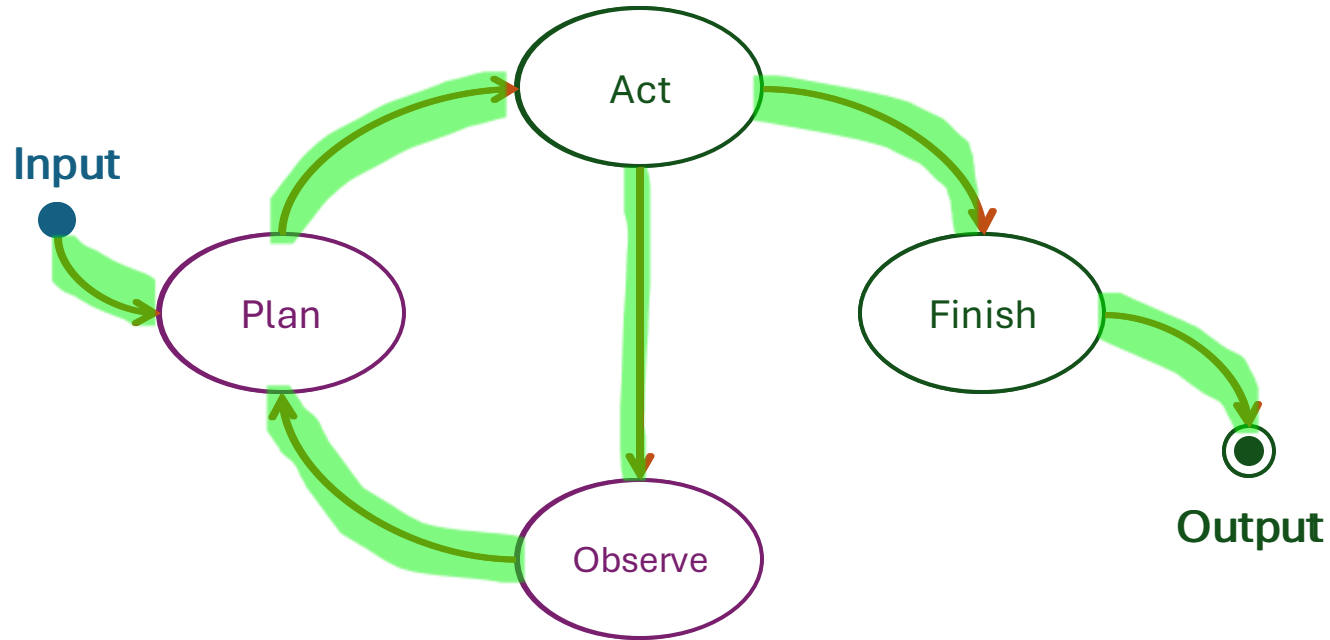
Object diagram





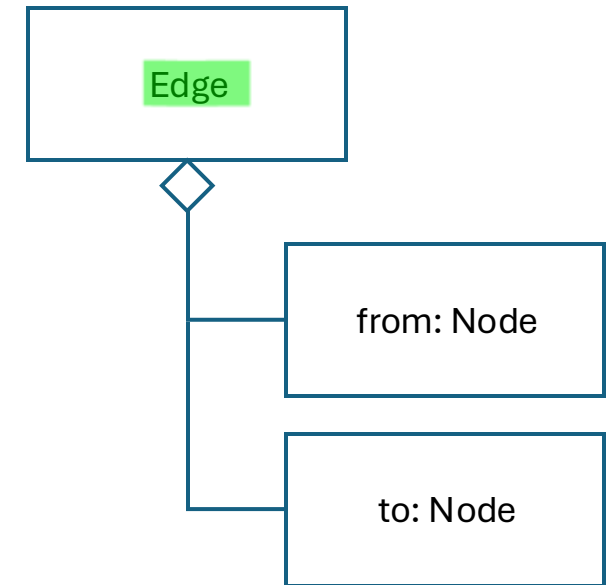
ReAct agent *graph prototype*

Formalization



- *class Edge*
 - Edge is a data structure for connecting two **Nodes**

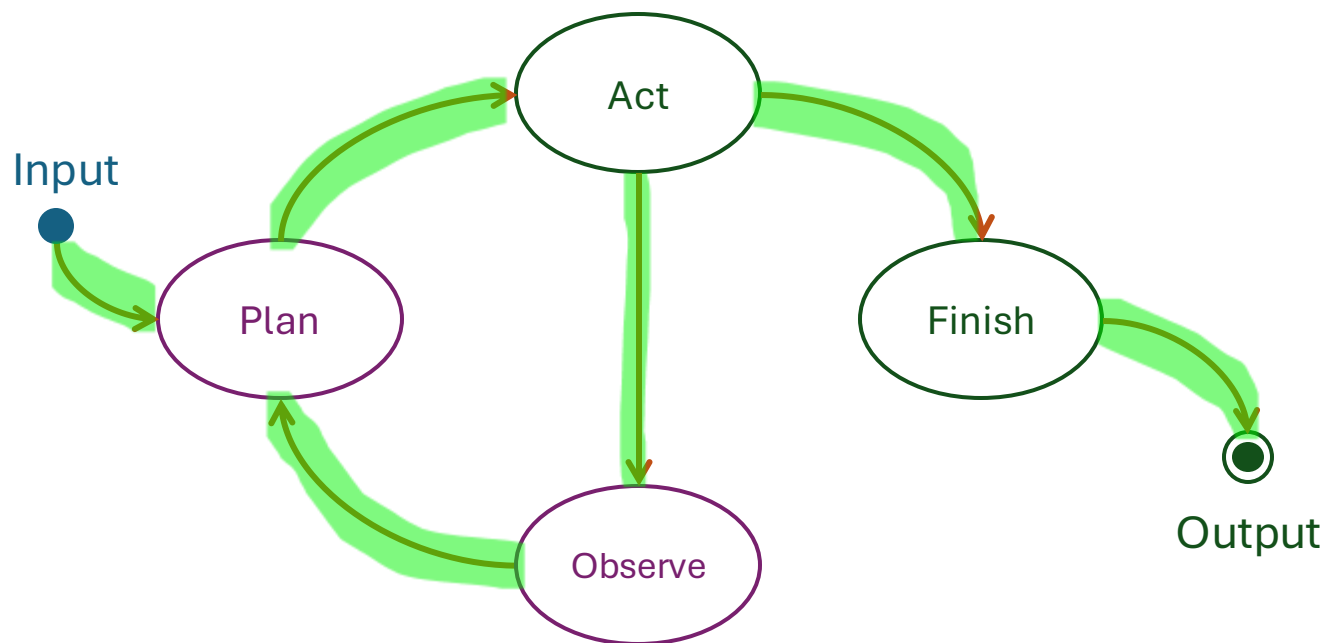
Class diagram





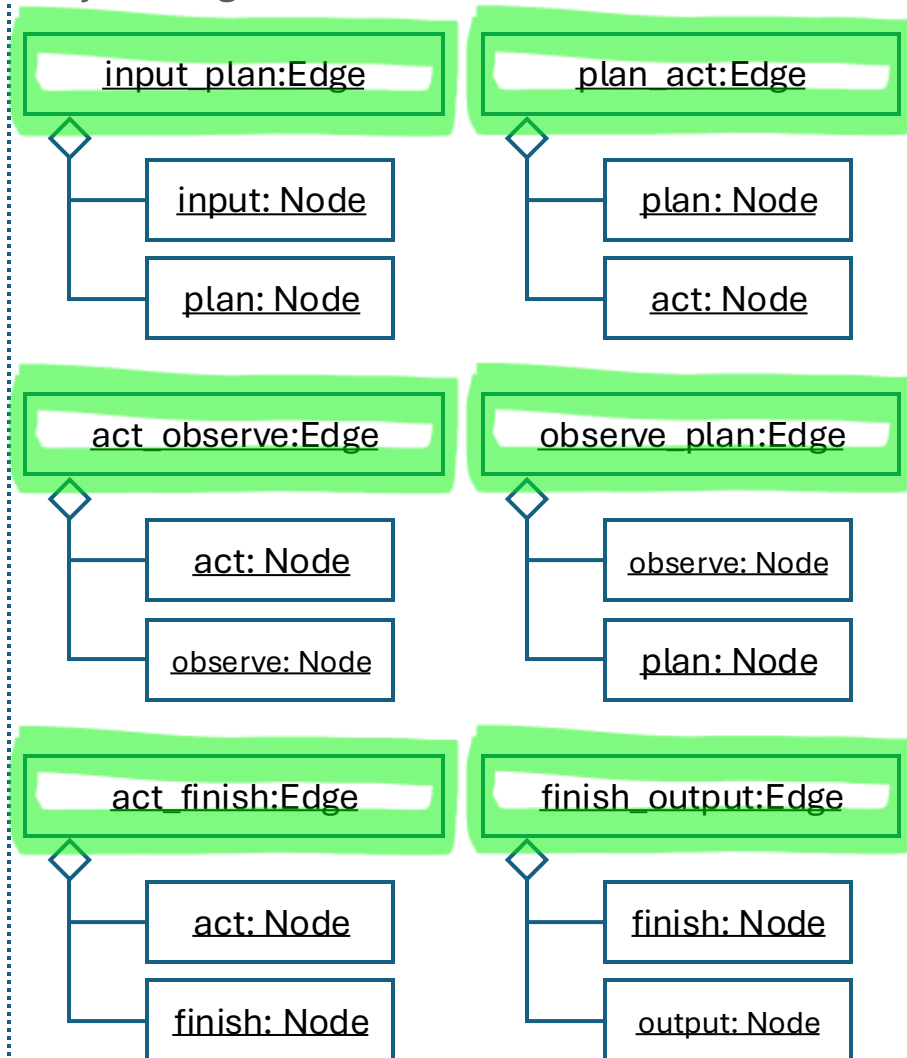
ReAct agent *graph prototype*

Formalization



- *class Edge*
 - Edge is a data structure for connecting two **Nodes**

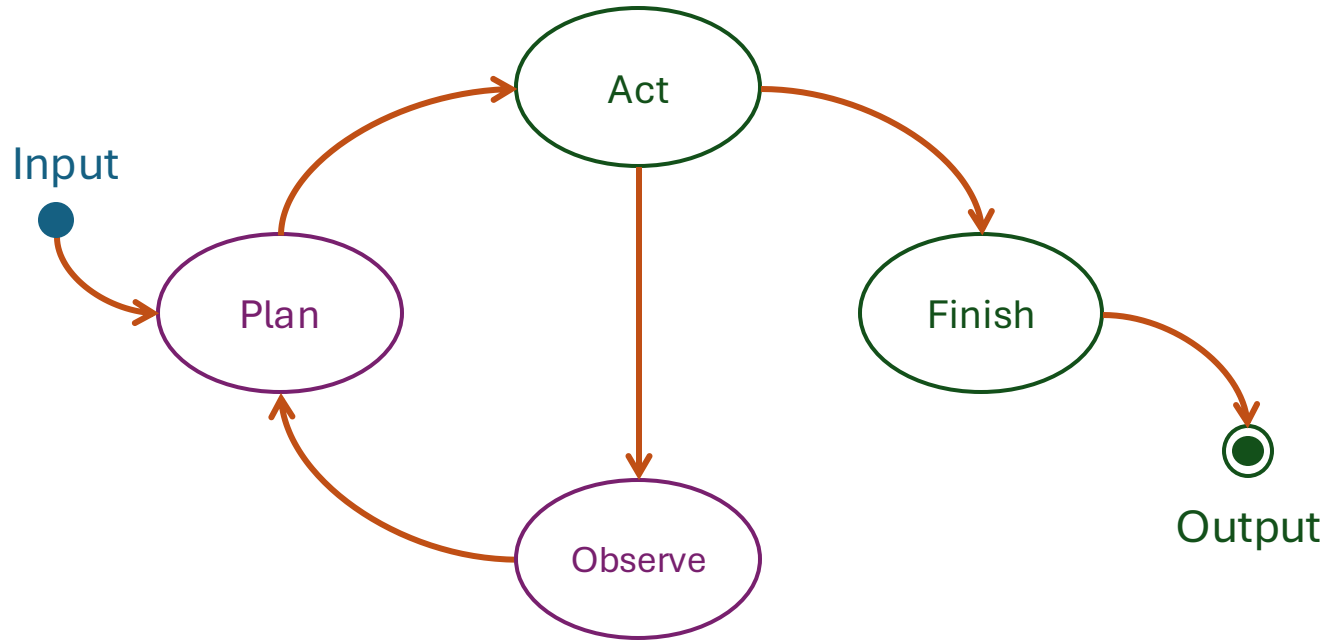
Object diagram





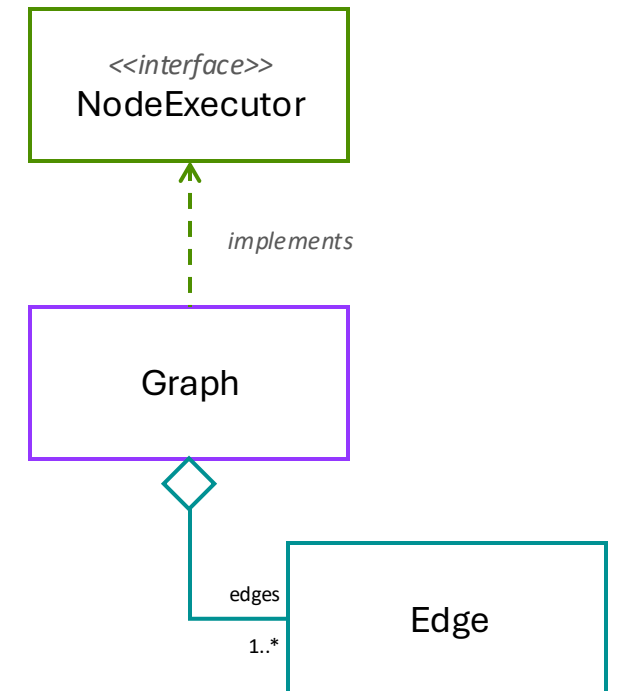
ReAct agent *graph* prototype

Formalization



- *class Graph*
 - Graph is the control class that traverse through node due to edges and executes their code.
 - Graph implements *NodeExecutor* so it can used as a regular node.

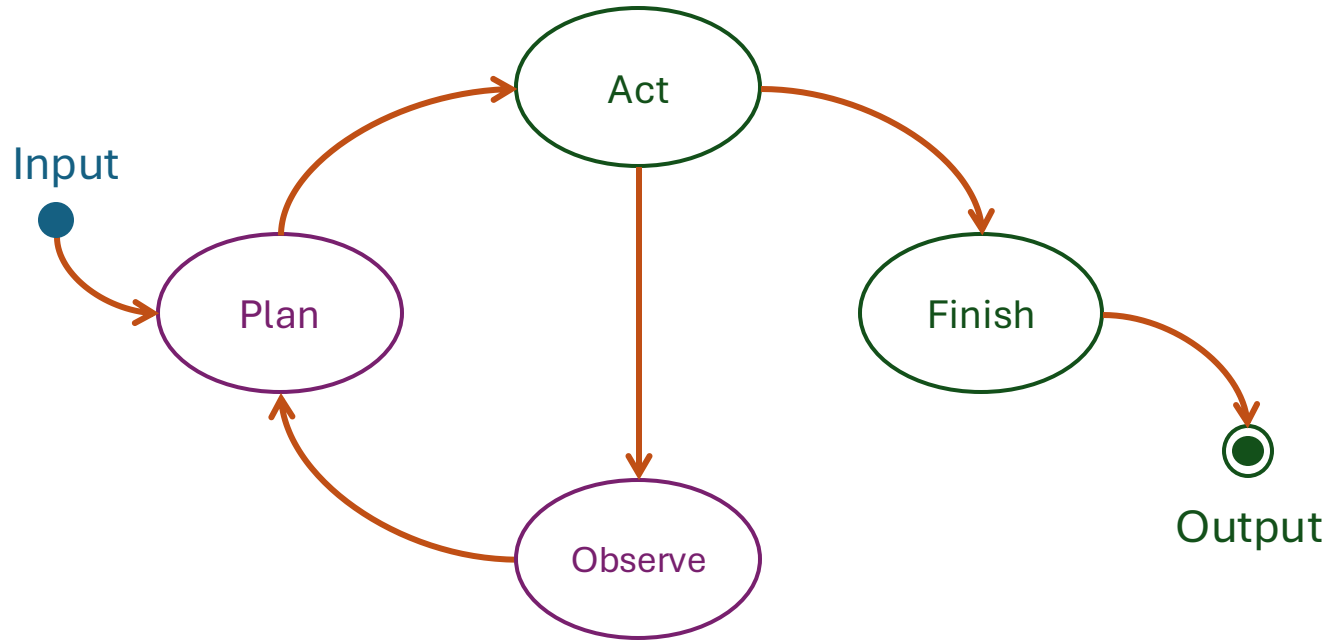
Class diagram





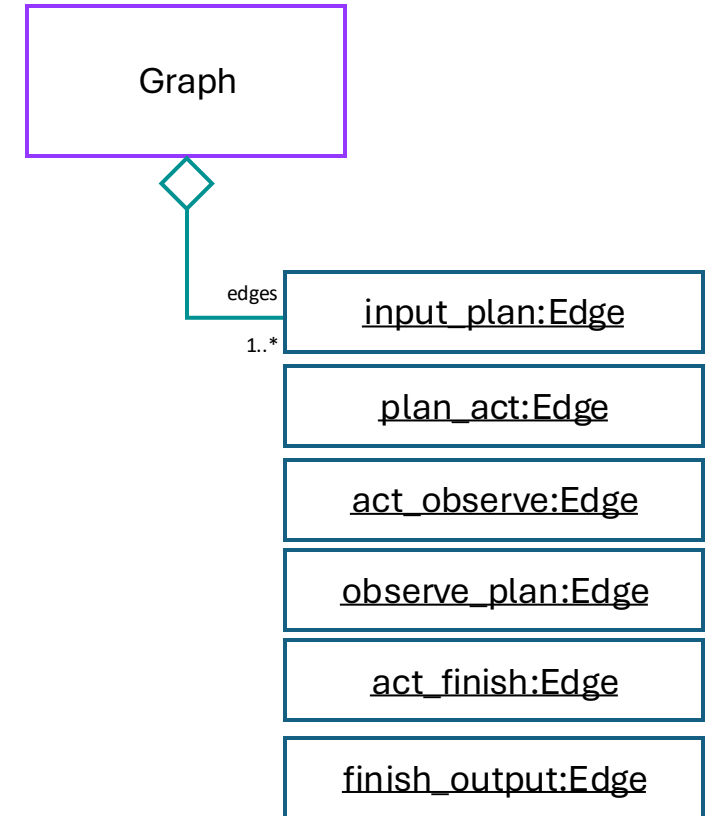
ReAct agent *graph prototype*

Formalization and prototype



- *class Graph*
 - Graph is the control class that traverse through node due to edges and executes their code.
 - Graph implements *NodeExecutor* so it can used as a regular node.

Object diagram





ReAct agent *graph prototype*

Formalization and prototype

- The prototype

- Url

- <https://github.ibm.com/Incubation/peri-agent-framework/pull/99>

- How to run

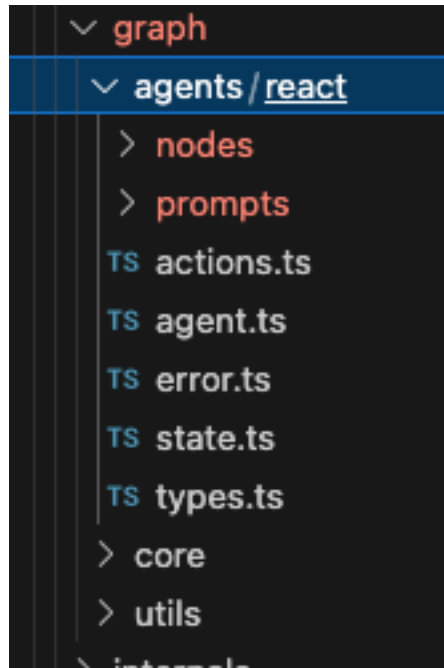
```
nvm use  
yarn  
yarn start:react:graph
```



ReAct agent *graph* prototype

Formalization and prototype

Graph package



```
You, 16 minutes ago | 1 author (You)
export class ReactAgent implements NodeExecutor<State> {
  private graph: Graph<State>;
  private availableActions: ActionDefinition[];
  public readonly logger: Logger;

  constructor({ modelId, tools, logger }: ReactAgentConfig) {
    this.logger = logger ?? Logger.of({ name: "graph.agent[react]" });
    this.logger.debug("ReAct[graph] agent has been created!");

    this.availableActions = getAvailableActions(tools);

    // Nodes
    const input = InputNode.getNode();
    const plan = PlanNode.getNode(modelId);
    const act = ActNode.getNode(tools);
    const observe = ObserveNode.getNode(modelId);
    const finish = FinishNode.getNode();
    const output = OutputNode.getNode();

    // Edges
    const edges = [
      new Edge(input, plan),
      new Edge(plan, act),
      new Edge(act, observe),
      new Edge(observe, plan),
      new Edge(act, finish),
      new Edge(finish, output),
    ];
    const start = edges[0];

    // Graph
    this.graph = new Graph(edges, start);

    execute(state: State): Promise<NodeExecutorOutput<State>> {
      return this.graph.execute({ ...state, availableActions: this.availableActions });
    }
  }
}
```

You, 11 minutes ago • feat(graph): add graph and implement simple rea...

ReAct agent setup



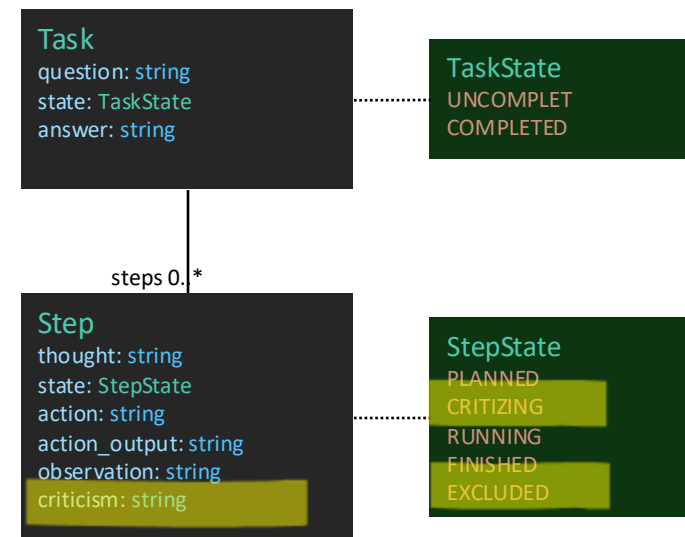
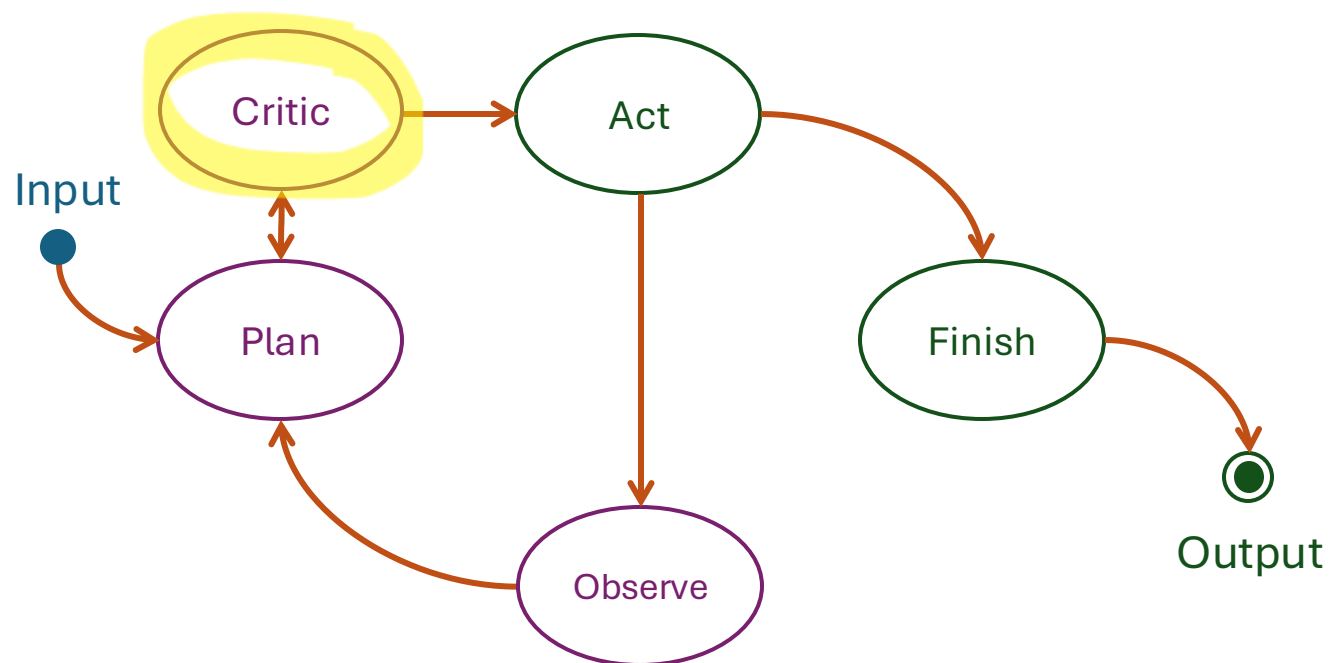
ReAct agent *evolution*

Critic



ReAct agent *graph prototype*

Critic



- Critic
 - An additional step behind plan where new steps are criticized
 - It is in loop with plan step
 - Technically: there is a new LLM call with specific prompt
- Motivation is to correct LLM mistakes and improve his performance



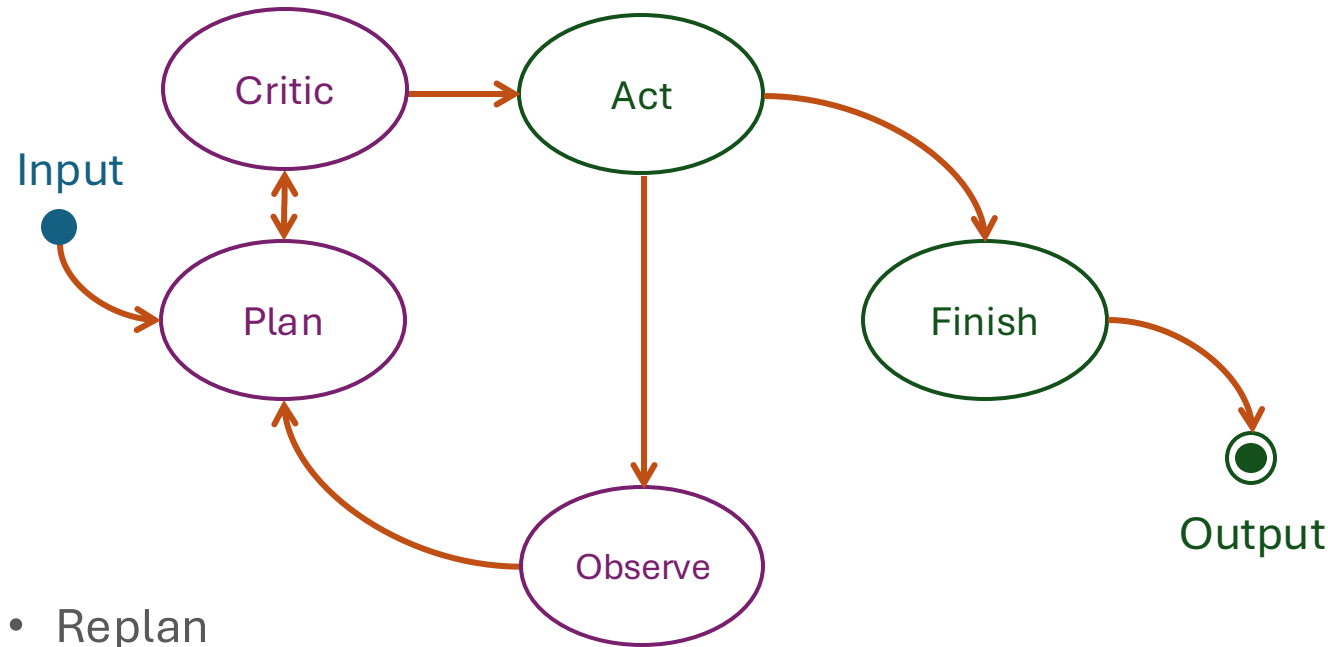
ReAct agent *evolution*

Critic + Replan

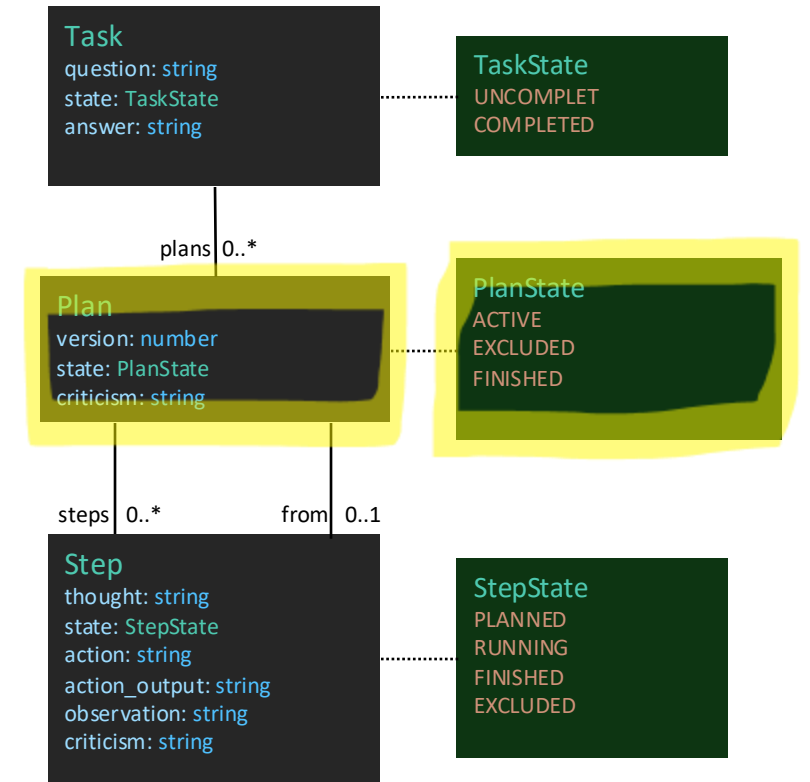


ReAct agent *evolution*

Critic + Replan



- Replan
 - Creates a plan trajectory with all steps immediately at the beginning
 - Plan trajectory is adapting
 - Technically: More complex state and logic in plan and critic steps
- Motivation
 - Recover from mistakes
 - Better traceability - user can see where the agent is going
 - Opening the possibility of how the user could create or change the agent's plan trajectory





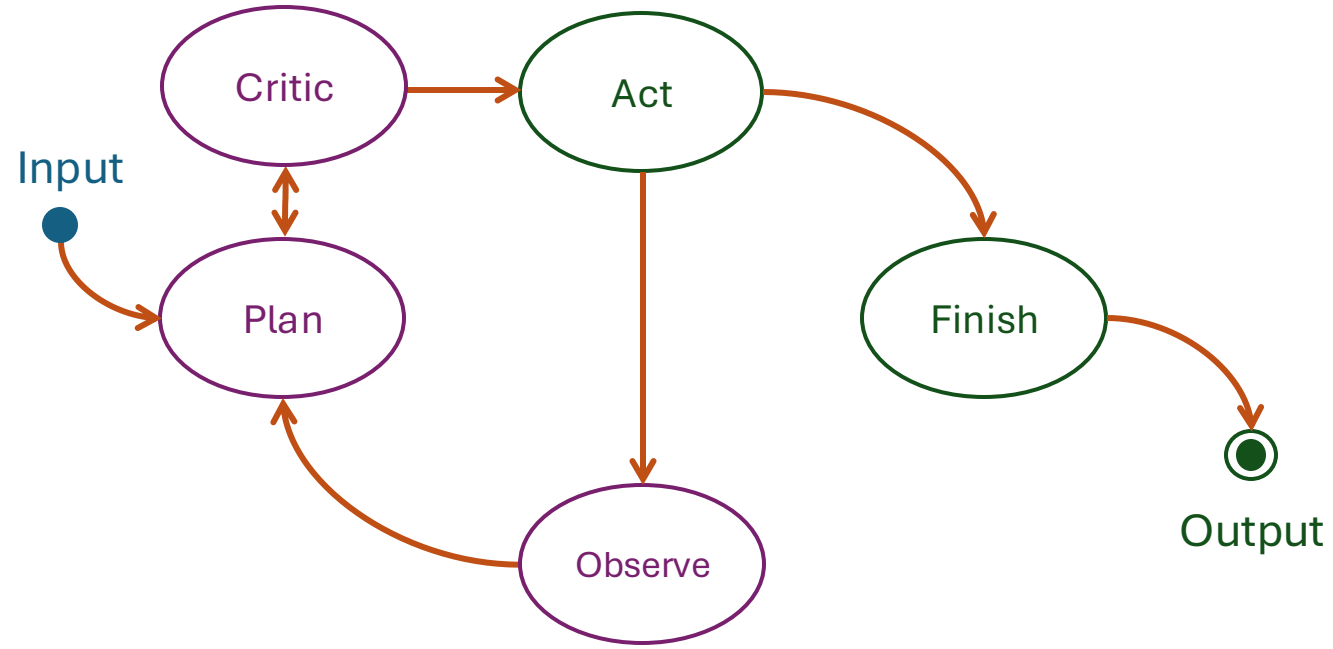
ReAct agent *evolution*

Critic + Replan + Human in the loop



ReAct agent *evolution*

Critic + Replan + Human in the loop

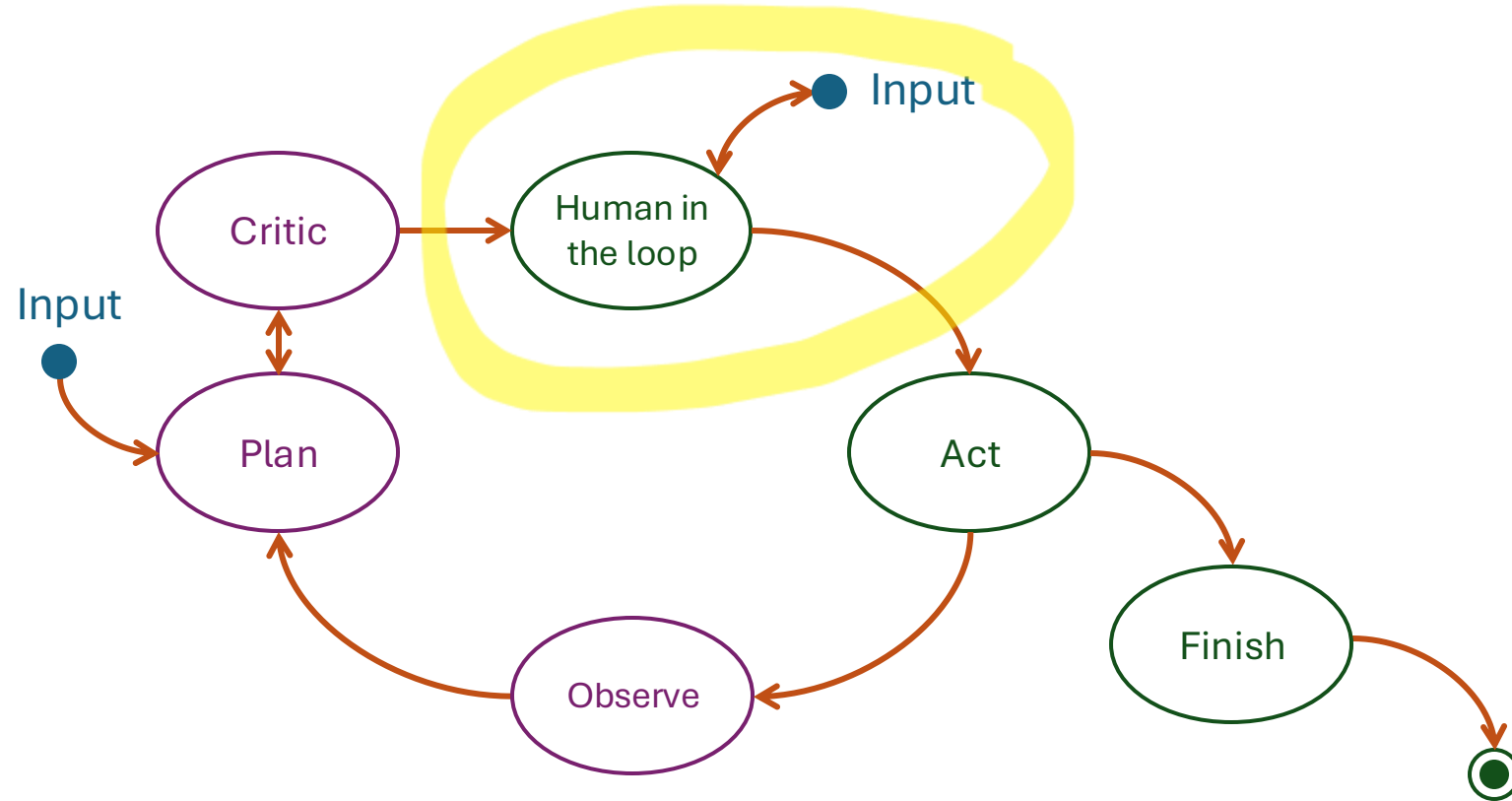


- Human in the loop



ReAct agent *evolution*

Critic + Replan + Human in the loop



- Human in the loop
 - User has a full control over the further actions
 - User can control an agent loop (▶ play / ⏸ pause / ■ stop / ◀◀ rewind)
 - Technically: An action like step that will allow to interrupt inner loop (state machine) and edit some parts of agent memory (state)



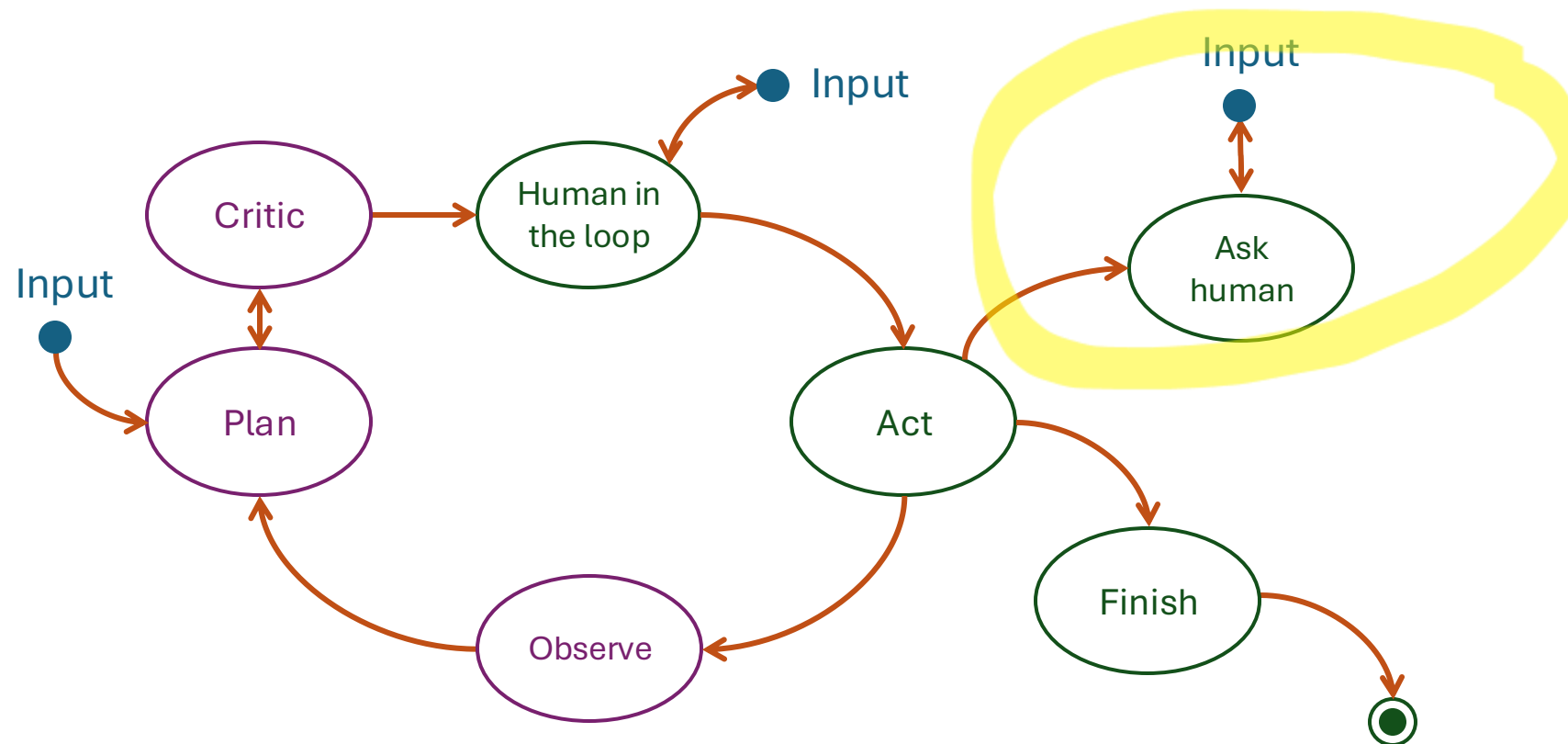
ReAct agent *evolution*

Critic + Replan + Human in the loop + Human in the action



ReAct agent *evolution*

Critic + Replan + Human in the loop + Human in the action



- Human in the action
 - User can be requested by agent like another tool
 - Technically: An action like step that can interact with user chat



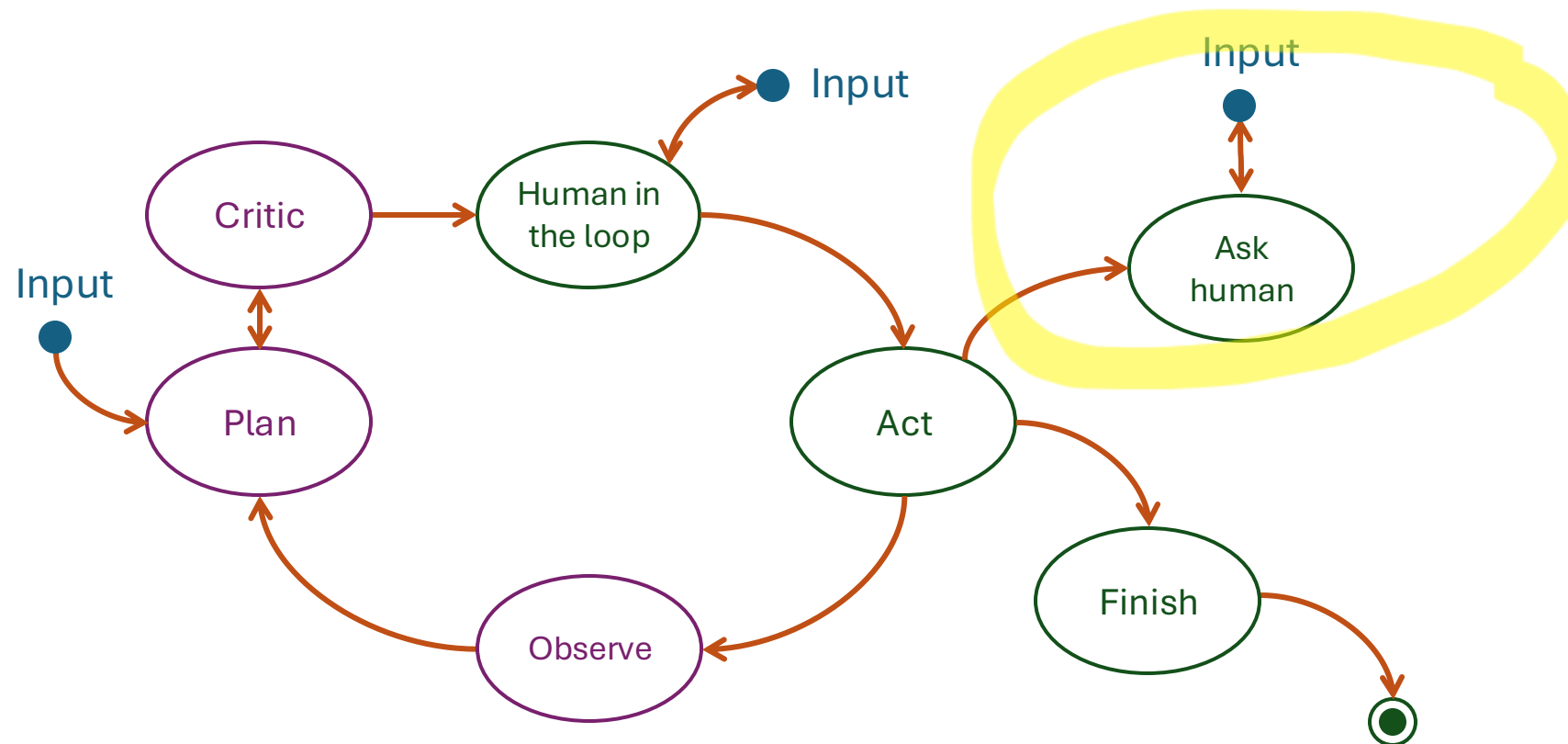
Suitable candidate for standalone agent specialized to get required input





ReAct agent *evolution*

Critic + Replan + Human in the loop + Human in the action



- Human in the action
 - User can be requested by agent like another tool
 - Technically: An action like step that can interact with user chat



Suitable candidate for standalone agent specialized to get required input





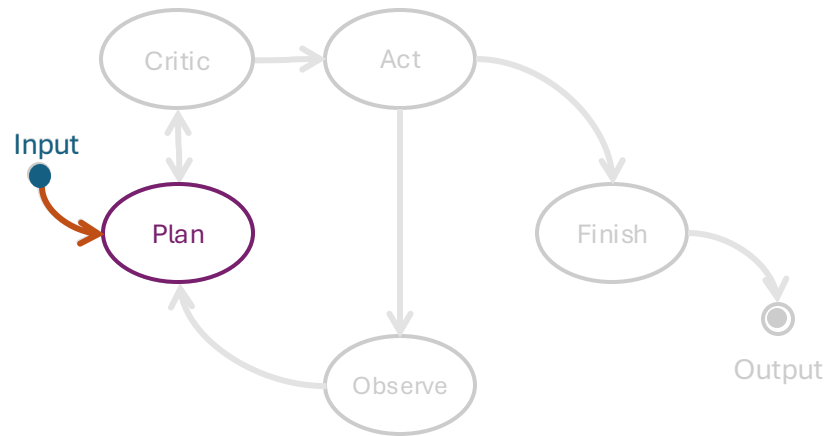
ReAct agent *evolution*

Critic + Replan *example*



ReAct agent *evolution*

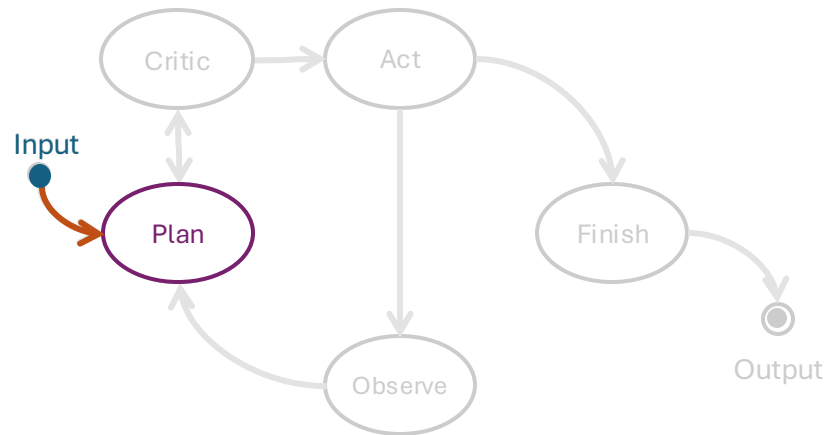
Critic + Replan *example*





ReAct agent *evolution*

Critic + Replan *example*



t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Active
criticism: null

s1:Step

thought: "I need to know the area of the UK."
state: Planned
action: "GeologicService[Area of the UK]"
action_output: null
observation: null
criticism: null

s2:Step

thought: "I need to know the area of the US."
state: Planned
action: "GeologicService[Area of the US]"
action_output: null
observation: null
criticism: null

s3:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

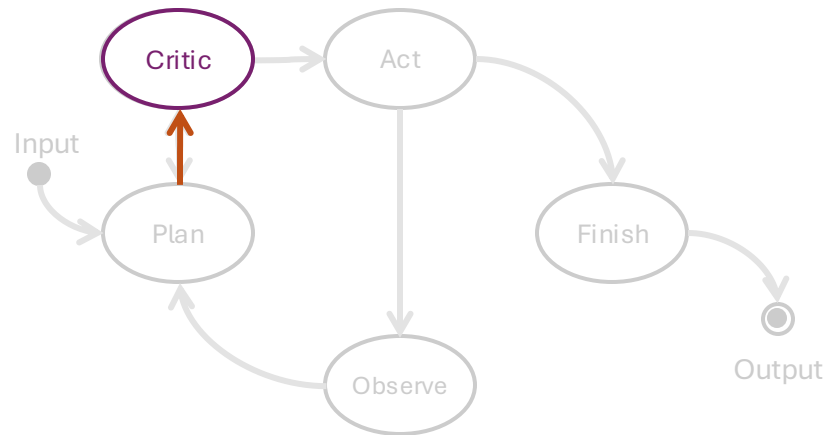
s4:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Active
criticism: null

s1:Step

thought: "I need to know the area of the UK."
state: Planned
action: "GeologicService[Area of the UK]"
action_output: null
observation: null
criticism: null

s2:Step

thought: "I need to know the area of the US."
state: Planned
action: "GeologicService[Area of the US]"
action_output: null
observation: null
criticism: null

s3:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

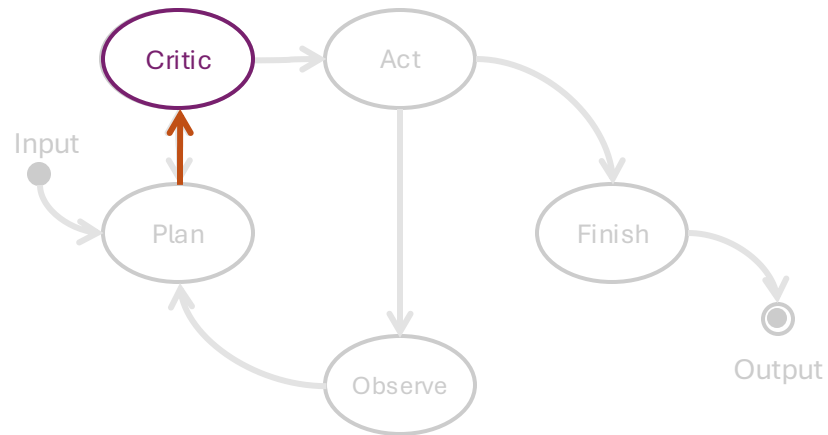
s4:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Active
criticism: null

s1:Step

thought: "I need to know the area of the UK."
state: Planned
action: "GeologicService[Area of the UK]"
action_output: null
observation: null
criticism: null

s2:Step

thought: "I need to know the area of the US."
state: Planned
action: "GeologicService[Area of the US]"
action_output: null
observation: null
criticism: null

s3:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

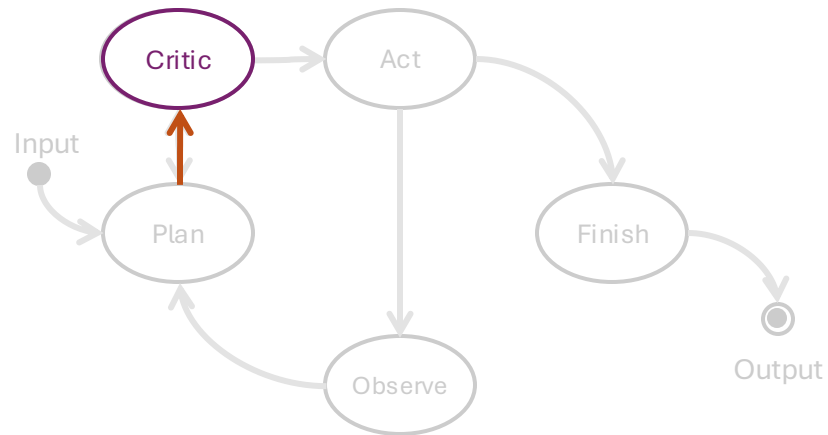
s4:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Active
criticism: null

s1:Step

thought: "I need to know the area of the UK."
state: Planned
action: "GeologicService[Area of the UK]"
action_output: null
observation: null
criticism: "GeologicService tool doesn't exists."

s2:Step

thought: "I need to know the area of the US."
state: Planned
action: "GeologicService[Area of the US]"
action_output: null
observation: null
criticism: null

s3:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

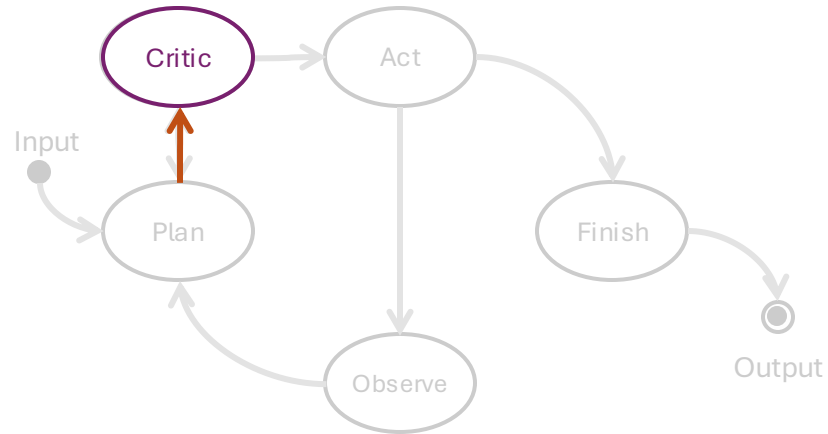
s4:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



*Plan mistakes
correction*

t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Active
criticism: null

s1:Step

thought: "I need to know the area of the UK."
state: Excluded
action: "GeologicService[Area of the UK]"
action_output: null
observation: null
criticism: "GeologicService tool doesn't exists."

s2:Step

thought: "I need to know the area of the US."
state: Planned
action: "GeologicService[Area of the US]"
action_output: null
observation: null
criticism: null

s3:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

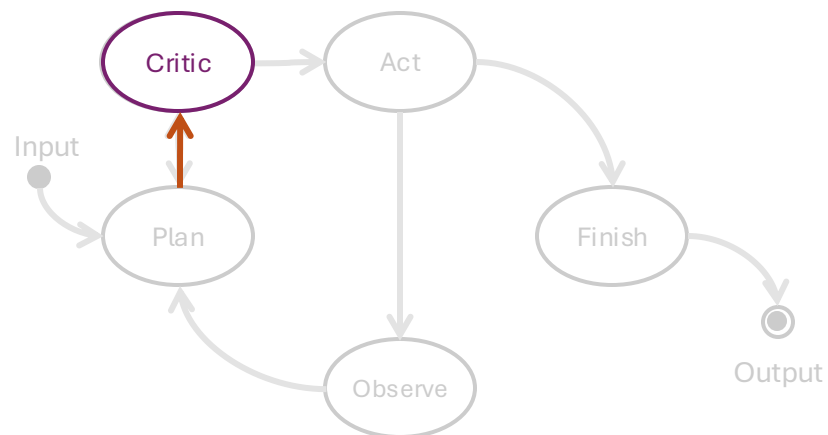
s4:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



Plan mistakes
correction

t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Active
criticism: null

s1:Step

thought: "I need to know the area of the UK."
state: Excluded
action: "GeologicService[Area of the UK]"
action_output: null
observation: null
criticism: "GeologicService tool doesn't exists."

s2:Step

thought: "I need to know the area of the US."
state: Planned
action: "GeologicService[Area of the US]"
action_output: null
observation: null
criticism: "GeologicService tool doesn't exists."

s3:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

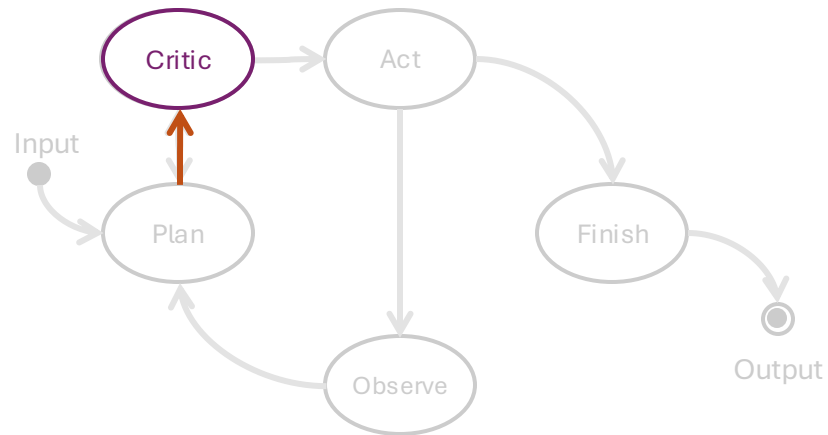
s4:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



*Plan mistakes
correction*

t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Active
criticism: null

s1:Step

thought: "I need to know the area of the UK."
state: Excluded
action: "GeologicService[Area of the UK]"
action_output: null
observation: null
criticism: "GeologicService tool doesn't exists."

s2:Step

thought: "I need to know the area of the US."
state: Excluded
action: "GeologicService[Area of the US]"
action_output: null
observation: null
criticism: "GeologicService tool doesn't exists."

s3:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

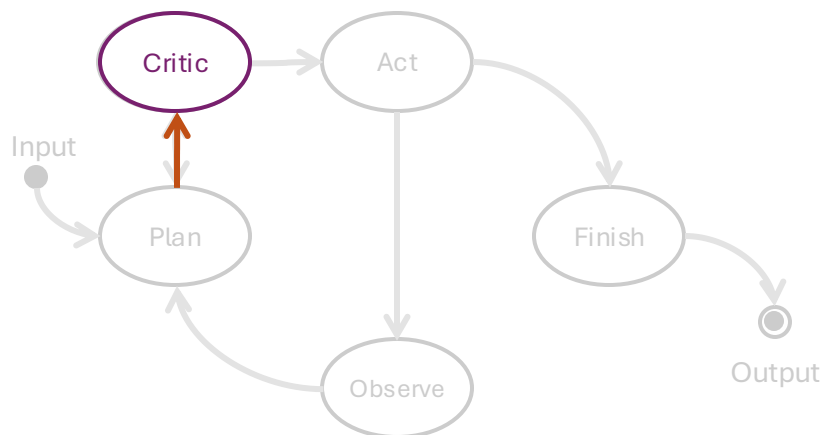
s4:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan example



Plan mistakes
correction

t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Active
criticism: "You can't use a GeologicService this action doesn't exist, use another action."

s1:Step

thought: "I need to know the area of the UK."
state: Excluded
action: "GeologicService[Area of the UK]"
action_output: null
observation: null
criticism: "GeologicService tool doesn't exists."

s2:Step

thought: "I need to know the area of the US."
state: Excluded
action: "GeologicService[Area of the US]"
action_output: null
observation: null
criticism: "GeologicService tool doesn't exists."

s3:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

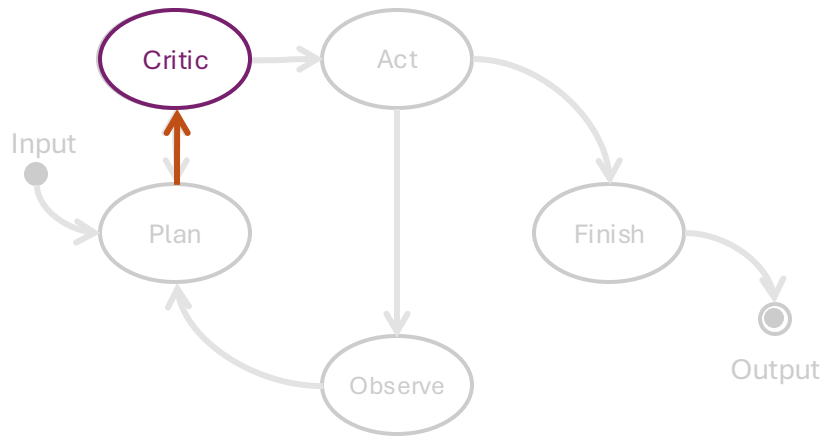
s4:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



Plan mistakes
correction

t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a GeologicService this action doesn't exist, use another action."

s1:Step

thought: "I need to know the area of the UK."
state: Excluded
action: "GeologicService[Area of the UK]"
action_output: null
observation: null
criticism: "GeologicService tool doesn't exists."

s2:Step

thought: "I need to know the area of the US."
state: Excluded
action: "GeologicService[Area of the US]"
action_output: null
observation: null
criticism: "GeologicService tool doesn't exists."

s3:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

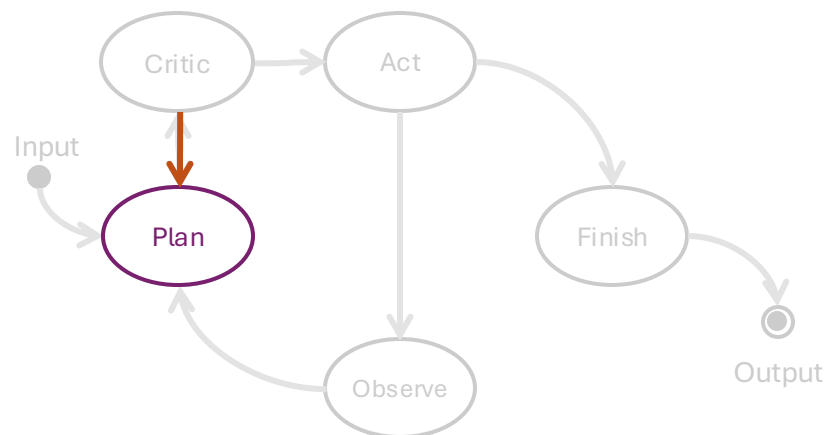
s4:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



*Plan mistakes
correction*

t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a C

p2:Plan

version: 2
state: Active
criticism: null

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the area of the UK."
state: Planned
action: "Wikipedia[Area of the UK]"
action_output: null
observation: null
criticism: null

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the area of the US."
state: Planned
action: "Wikipedia[Area of the US]"
action_output: null
observation: null
criticism: null

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

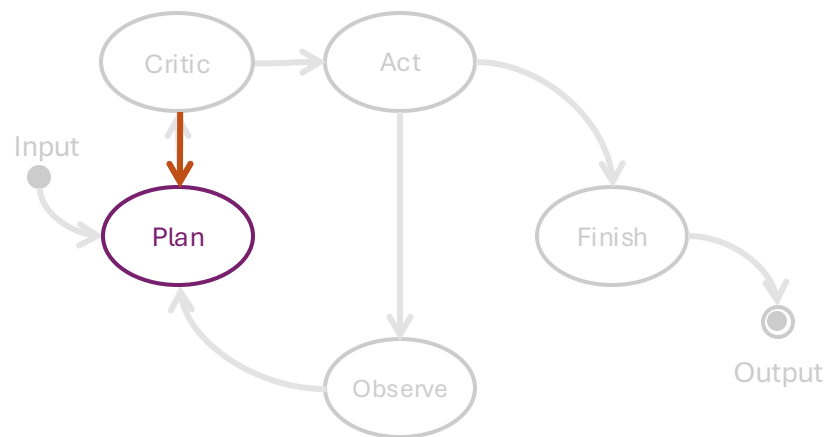
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



*Plan mistakes
correction*

t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a C

p2:Plan

version: 2
state: Active
criticism: null

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the area of the UK."
state: Planned
action: "Wikipedia[Area of the UK]"
action_output: null
observation: null
criticism: null

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the area of the US."
state: Planned
action: "Wikipedia[Area of the US]"
action_output: null
observation: null
criticism: null

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

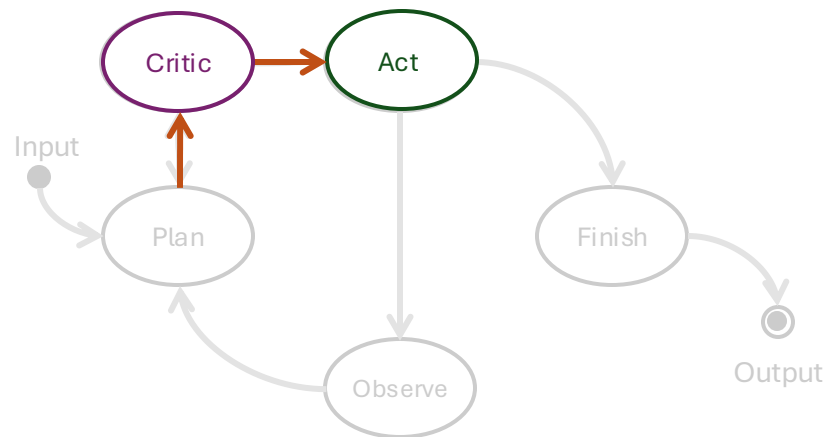
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a C

p2:Plan

version: 2
state: Active
criticism: null

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the area of the UK."
state: Running
action: "Wikipedia[Area of the UK]"
action_output: null
observation: null
criticism: null

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the area of the US."
state: Planned
action: "Wikipedia[Area of the US]"
action_output: null
observation: null
criticism: null

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

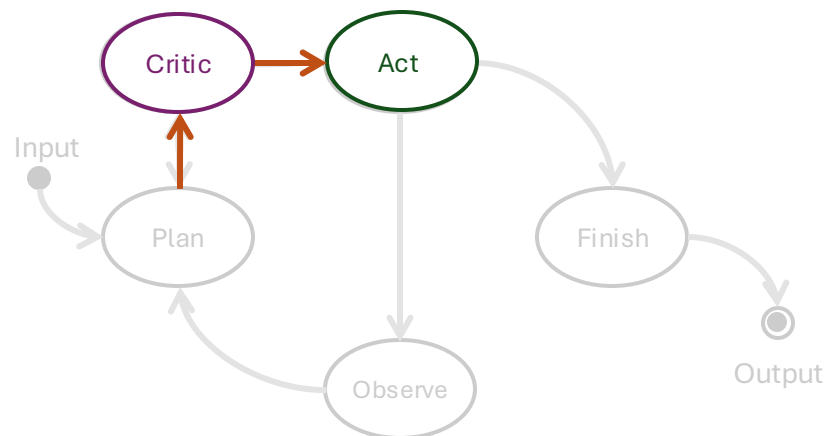
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a C

p2:Plan

version: 2
state: Active
criticism: null

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the area of the UK."
state: Running
action: "Wikipedia[Area of the UK]"
action_output: "The United Kingdom of Great Britain and Northern Ireland..."
observation: null
criticism: null

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the area of the US."
state: Planned
action: "Wikipedia[Area of the US]"
action_output: null
observation: null
criticism: null

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

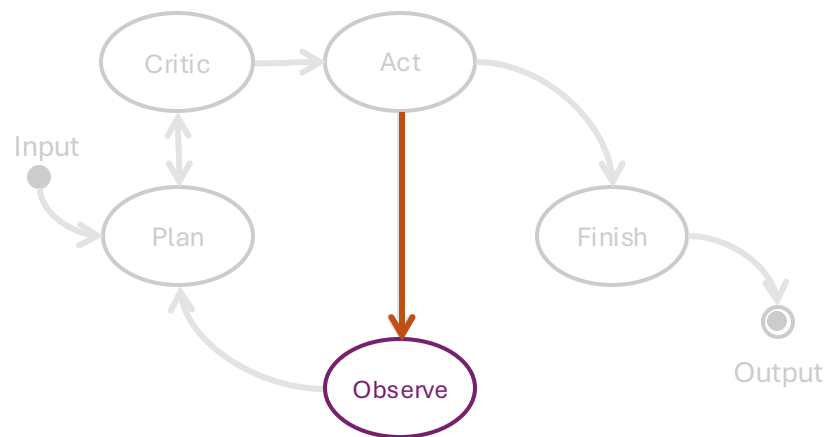
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a C

p2:Plan

version: 2
state: Active
criticism: null

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the area of the UK."
state: Running
action: "Wikipedia[Area of the UK]"
action_output: "The United Kingdom of Great Britain and Northern Ireland..."
observation: null
criticism: null

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the area of the US."
state: Planned
action: "Wikipedia[Area of the US]"
action_output: null
observation: null
criticism: null

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

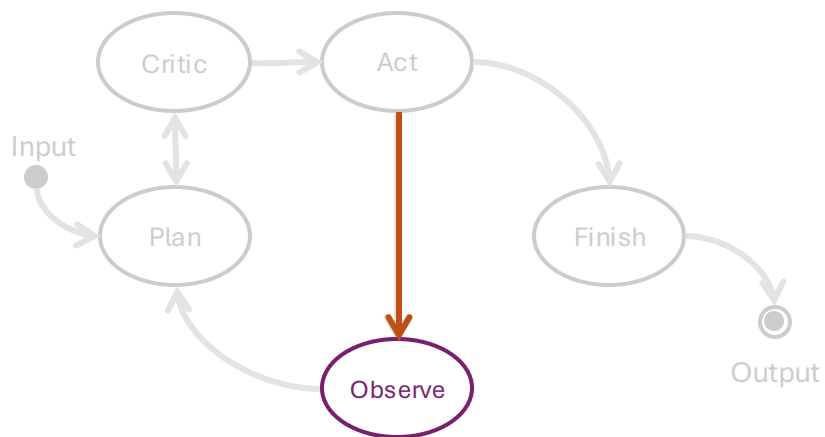
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a C

p2:Plan

version: 2
state: Active
criticism: null

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the area of the UK."
state: Finished
action: "Wikipedia[Area of the UK]"
action_output: "The United Kingdom of Great Britain and Northern Ireland..."
observation: "UK area is 244,376 km2"
criticism: null

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the area of the US."
state: Planned
action: "Wikipedia[Area of the US]"
action_output: null
observation: null
criticism: null

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

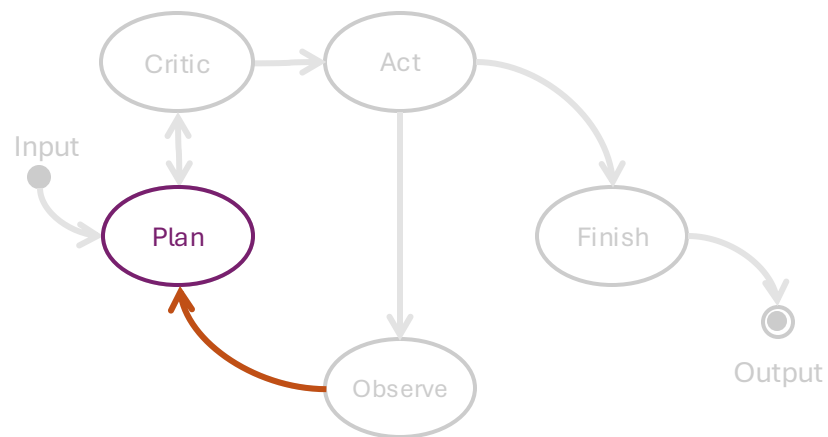
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a C

p2:Plan

version: 2
state: Active
criticism: null

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the area of the UK."
state: Finished
action: "Wikipedia[Area of the UK]"
action_output: "The United Kingdom of Great Britain and Northern Ireland..."
observation: "UK area is 244,376 km2"
criticism: null

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the area of the US."
state: Planned
action: "Wikipedia[Area of the US]"
action_output: null
observation: null
criticism: null

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

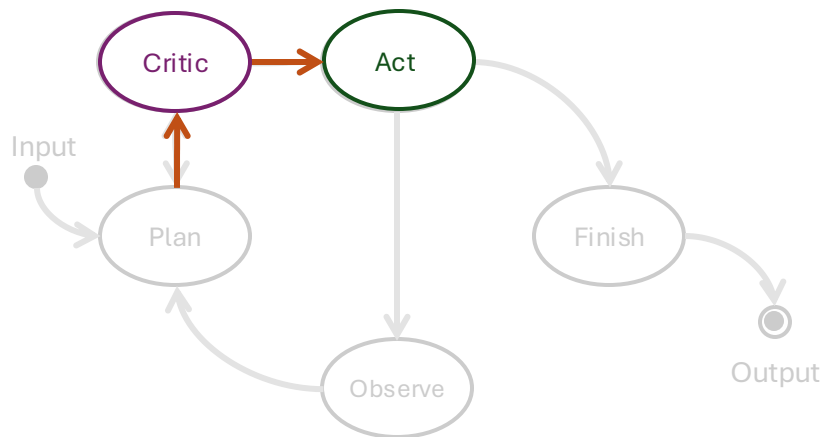
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a C

p2:Plan

version: 2
state: Active
criticism: null

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the area of the UK."
state: Finished
action: "Wikipedia[Area of the UK]"
action_output: "The United Kingdom of Great Britain and Northern Ireland..."
observation: "UK area is 244,376 km2"
criticism: null

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the area of the US."
state: Running
action: "Wikipedia[Area of the US]"
action_output: "The United States of America (USA or U.S.A.), commonly known as..."
observation: null
criticism: null

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

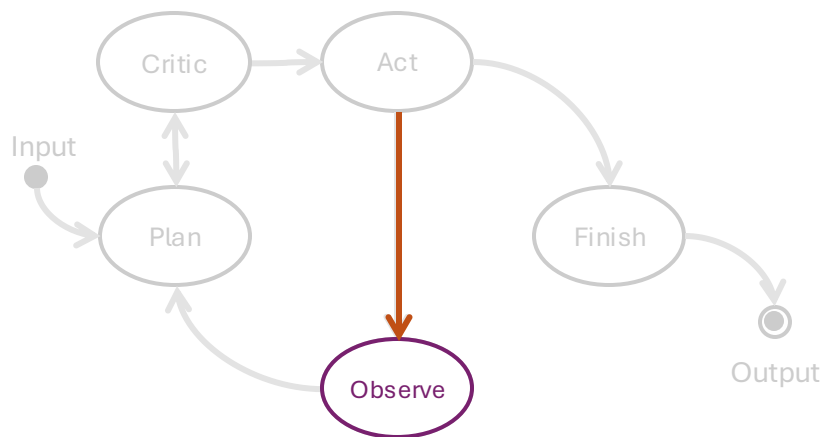
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a C

p2:Plan

version: 2
state: Active
criticism: null

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the area of the UK."
state: Finished
action: "Wikipedia[Area of the UK]"
action_output: "The United Kingdom of Great Britain and Northern Ireland..."
observation: "UK area is 244,376 km2"
criticism: null

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the area of the US."
state: Running
action: "Wikipedia[Area of the US]"
action_output: "The United States of America (USA or U.S.A.), commonly known as..."
observation: null
criticism: null

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

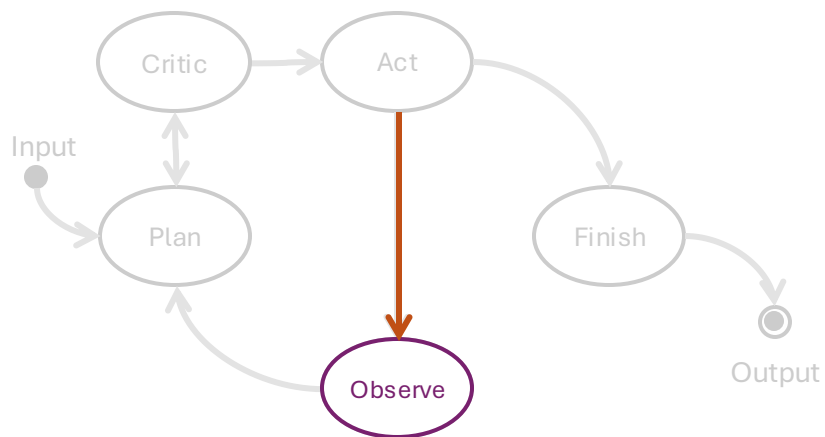
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



Tool output
reflexion

t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a C

p2:Plan

version: 2
state: Active
criticism: null

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the area of the UK."
state: Finished
action: "Wikipedia[Area of the UK]"
action_output: "The United Kingdom of Great Britain and Northern Ireland..."
observation: "UK area is 244,376 km2"
criticism: null

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the area of the US."
state: Running
action: "Wikipedia[Area of the US]"
action_output: "The United States of America (USA or U.S.A.), commonly known as..."
observation: "There is no are of the US"
criticism: null

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

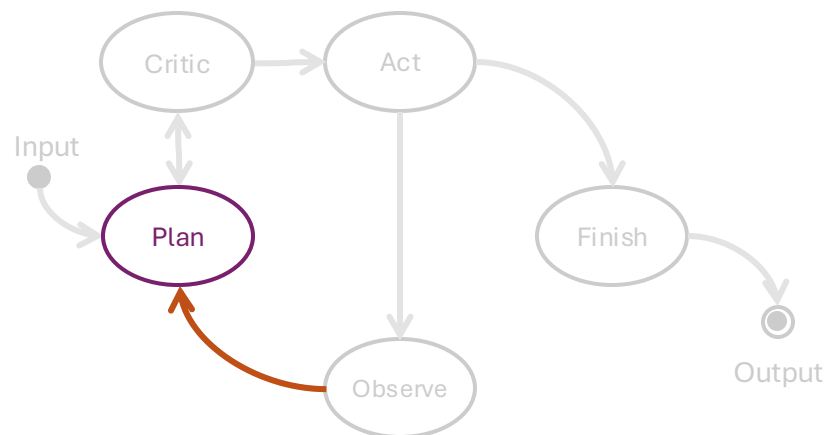
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



Tool output
reflexion

t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a C

p2:Plan

version: 2
state: Active
criticism: null

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the area of the UK."
state: Finished
action: "Wikipedia[Area of the UK]"
action_output: "The United Kingdom of Great Britain and Northern Ireland..."
observation: "UK area is 244,376 km2"
criticism: null

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the area of the US."
state: Finished
action: "Wikipedia[Area of the US]"
action_output: "The United States of America (USA or U.S.A.), commonly known as..."
observation: "There is no are of the US"
criticism: null

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

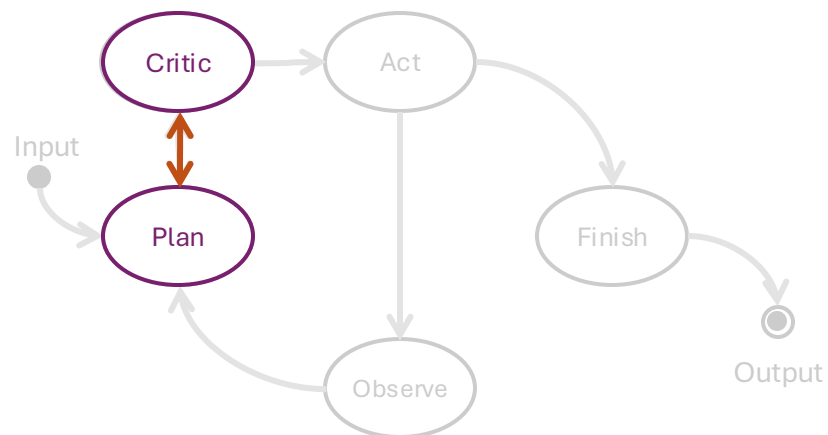
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan example



Tool output
reflexion

t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a C

p2:Plan

version: 2
state: Active
criticism: null

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the area of the UK."
state: Finished
action: "Wikipedia[Area of the UK]"
action_output: "The United Kingdom of Great Britain and Northern Ireland..."
observation: "UK area is 244,376 km2"
criticism: null

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the area of the US."
state: Finished
action: "Wikipedia[Area of the US]"
action_output: "The United States of America (USA or U.S.A.), commonly known as..."
observation: "There is no are of the US"
criticism: null

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

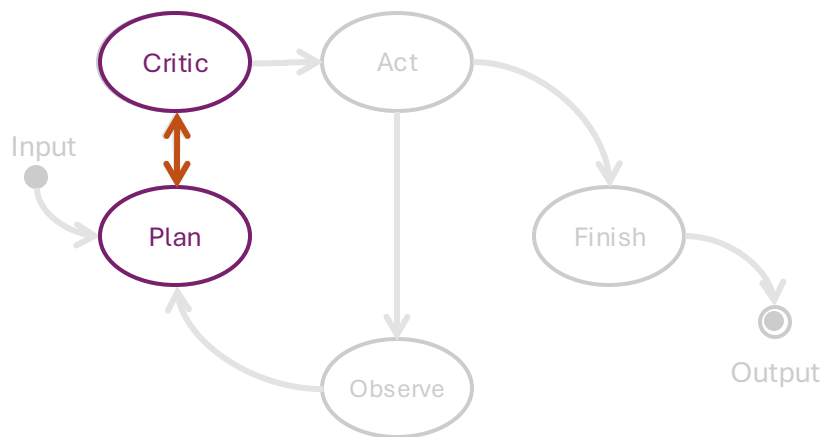
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



Tool output
reflexion

t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a C

p2:Plan

version: 2
state: Active
criticism: null

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the area of the UK."
state: Finished
action: "Wikipedia[Area of the UK]"
action_output: "The United Kingdom of Great Britain and Northern Ireland..."
observation: "UK area is 244,376 km2"
criticism: null

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the area of the US."
state: Finished
action: "Wikipedia[Area of the US]"
action_output: "The United States of America (USA or U.S.A.), commonly known as..."
observation: "There is no are of the US"
criticism: "Step doesn't contain required value"

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

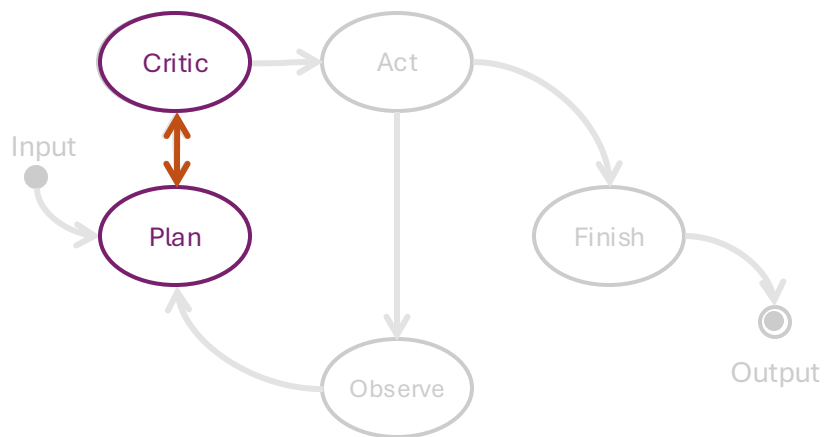
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



Tool output
reflexion

t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a C

p2:Plan

version: 2
state: Active
criticism: null

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the area of the UK."
state: Finished
action: "Wikipedia[Area of the UK]"
action_output: "The United Kingdom of Great Britain and Northern Ireland..."
observation: "UK area is 244,376 km2"
criticism: null

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the area of the US."
state: Excluded
action: "Wikipedia[Area of the US]"
action_output: "The United States of America (USA or U.S.A.), commonly known as..."
observation: "There is no are of the US"
criticism: "Step doesn't contain required value"

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

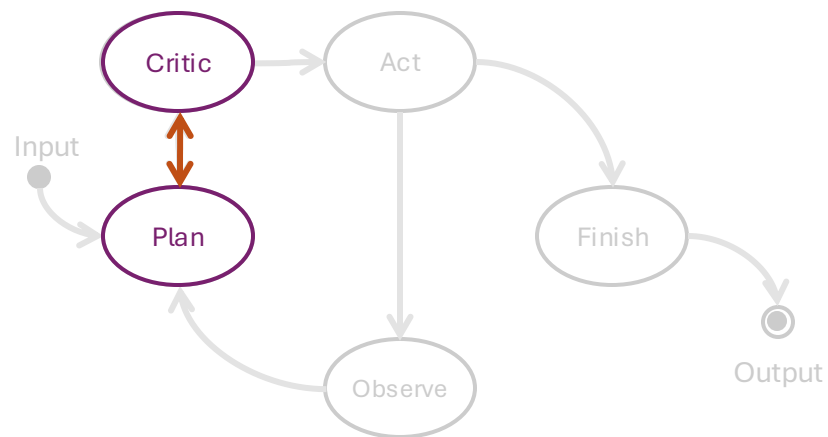
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan example



Tool output
reflexion

t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a G

p2:Plan

version: 2
state: Active
criticism: "Step for getting area of US didn't produced any useful value."

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the area of the UK."
state: Finished
action: "Wikipedia[Area of the UK]"
action_output: "The United Kingdom of Great Britain and Northern Ireland..."
observation: "UK area is 244,376 km2"
criticism: null

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the area of the US."
state: Excluded
action: "Wikipedia[Area of the US]"
action_output: "The United States of America (USA or U.S.A.), commonly known as..."
observation: "There is no are of the US"
criticism: "Step doesn't contain required value"

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

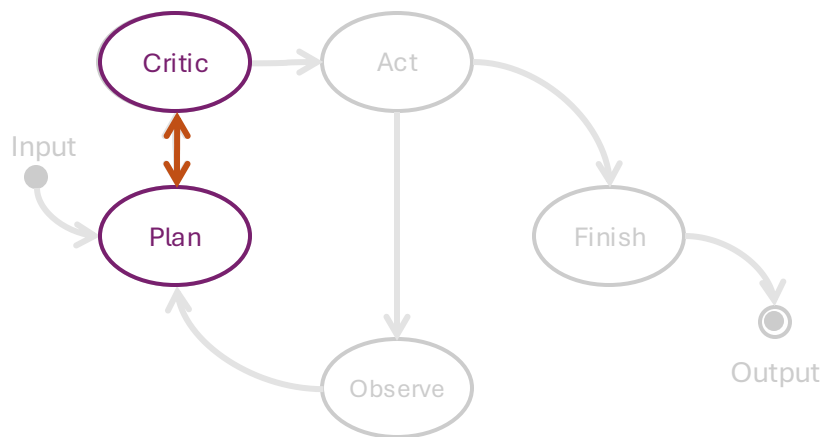
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan example



Tool output
reflexion

t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a G

p2:Plan

version: 2
state: Excluded
criticism: "Step for getting area of US didn't produced any useful value."

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the area of the UK."
state: Finished
action: "Wikipedia[Area of the UK]"
action_output: "The United Kingdom of Great Britain and Northern Ireland..."
observation: "UK area is 244,376 km2"
criticism: null

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the area of the US."
state: Excluded
action: "Wikipedia[Area of the US]"
action_output: "The United States of America (USA or U.S.A.), commonly known as..."
observation: "There is no are of the US"
criticism: "Step doesn't contain required value"

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

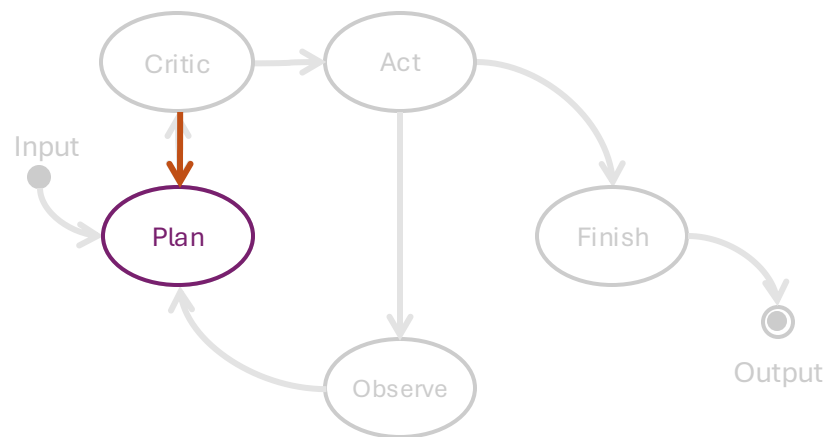
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



Tool output
reflexion

t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a G

p2:Plan

version: 2
state: Excluded
criticism: "Step for getting area of US didn't produced any useful value."

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the area of the UK."
state: Finished
action: "Wikipedia[Area of the UK]"
action_output: "The United Kingdom of Great Britain and Northern Ireland..."
observation: "UK area is 244,376 km2"
criticism: null

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Ar
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the area of the US."
state: Excluded
action: "Wikipedia[Area of the US]"
action_output: "The United States of America (USA or U.S.A.), commonly known as..."
observation: "There is no are of the US"
criticism: "Step doesn't contain required value"

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

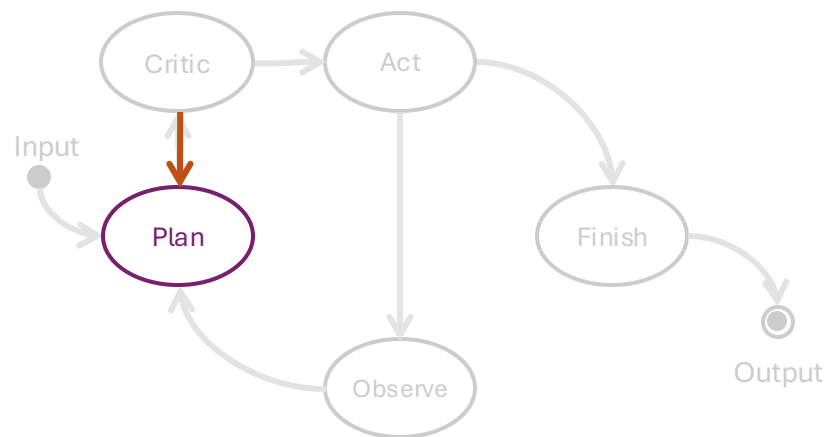
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



Tool output
reflexion

t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a G

p2:Plan

version: 2
state: Excluded
criticism: "Step for getting a

p3:Plan

version: 3
state: Active
criticism: null
from: s6:Step

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the
state: Finished
action: "Wikipedia[Area of t
action_output: "The United
observation: "UK area is 24
criticism: null

s9:Step

thought: "I need to know the area of the US."
state: Planned
action: "Google[Area of the US]"
action_output: null
observation: null

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the area of the US."
state: Excluded
action: "Wikipedia[Area of the US]"
action_output: "The United States of America (USA or US)"
observation: "There is no are of the US"
criticism: "Step doesn't contain required value"

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both areas I can calculate how many
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

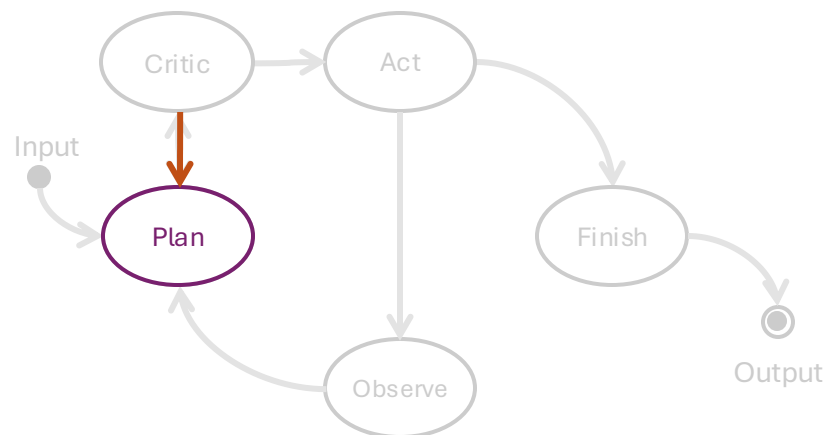
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



Tool output
reflexion

t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a G

p2:Plan

version: 2
state: Excluded
criticism: "Step for getting a

p3:Plan

version: 3
state: Active
criticism: null
from: s6:Step

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the
state: Finished
action: "Wikipedia[Area of t
action_output: "The United
observation: "UK area is 24
criticism: null

s9:Step

thought: "I need to know the area of the US."
state: Planned
action: "Google[Area of the US]"
action_output: null
observation: null

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the area of the US."
state: Excluded
action: "Wikipedia[Area of the US]"
action_output: "The United States of America (USA or US)
observation: "There is no are of the US"
criticism: "Step doesn't contain required value"

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both areas I can calculate how many
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

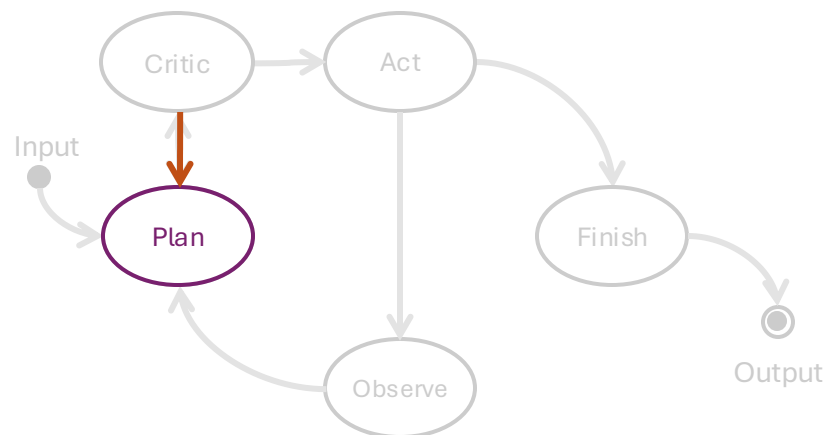
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



Tool output
reflexion

t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a G

p2:Plan

version: 2
state: Excluded
criticism: "Step for getting a

p3:Plan

version: 3
state: Active
criticism: null
from: s6:Step

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Area of t
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the
state: Finished
action: "Wikipedia[Area of t
action_output: "The United
observation: "UK area is 24
criticism: null

s9:Step

thought: "I need to know the area of the US."
state: Planned
action: "Google[Area of the US]"
action_output: null
observation: null

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Area of t
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the
state: Excluded
action: "Wikipedia[Area of t
action_output: "The United
observation: "There is no ar
criticism: "Step doesn't cont

s10:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating t
action_output: null
observation: null

s7:Step

thought: "When I have both
state: Planned
action: "Python[Calculating t
action_output: null
observation: null

s11:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

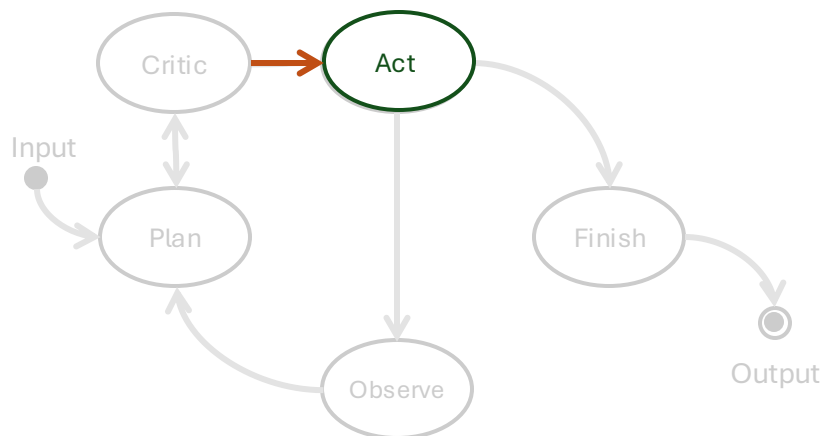
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a G

p2:Plan

version: 2
state: Excluded
criticism: "Step for getting a

p3:Plan

version: 3
state: Active
criticism: null
from: s6:Step

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Area of t
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the
state: Finished
action: "Wikipedia[Area of t
action_output: "The United
observation: "UK area is 24
criticism: null

s9:Step

thought: "I need to know the area of the US."
state: Planned
action: "Google[Area of the US]"
action_output: null
observation: null

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Area of t
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the
state: Excluded
action: "Wikipedia[Area of t
action_output: "The United
observation: "There is no ar
criticism: "Step doesn't cont

s10:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating t
action_output: null
observation: null

s7:Step

thought: "When I have both
state: Planned
action: "Python[Calculating t
action_output: null
observation: null

s11:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

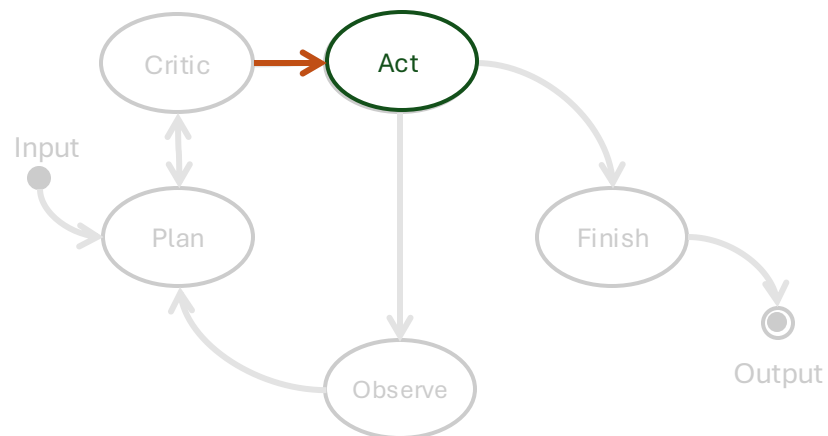
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a G

p2:Plan

version: 2
state: Excluded
criticism: "Step for getting a

p3:Plan

version: 3
state: Active
criticism: null
from: s6:Step

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the
state: Finished
action: "Wikipedia[Area of t
action_output: "The United
observation: "UK area is 24
criticism: null

s9:Step

thought: "I need to know the area of the US."
state: Running
action: "Google[Area of the US]"
action_output: null
observation: null

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the
state: Excluded
action: "Wikipedia[Area of t
action_output: "The United
observation: "There is no ar
criticism: "Step doesn't cont

s10:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s11:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

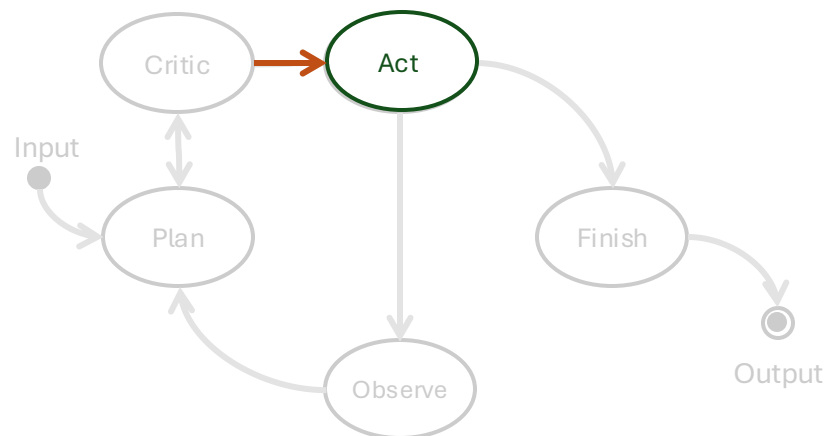
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a G

p2:Plan

version: 2
state: Excluded
criticism: "Step for getting a

p3:Plan

version: 3
state: Active
criticism: null
from: s6:Step

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the
state: Finished
action: "Wikipedia[Area of t
action_output: "The United
observation: "UK area is 24
criticism: null

s9:Step

thought: "I need to know the area of the US."
state: Running
action: "Google[Area of the US]"
action_output: "The area of the US is 3,809,525 square miles (9,840,000 square km)"
observation: null

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the
state: Excluded
action: "Wikipedia[Area of t
action_output: "The United
observation: "There is no ar
criticism: "Step doesn't cont

s10:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s11:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

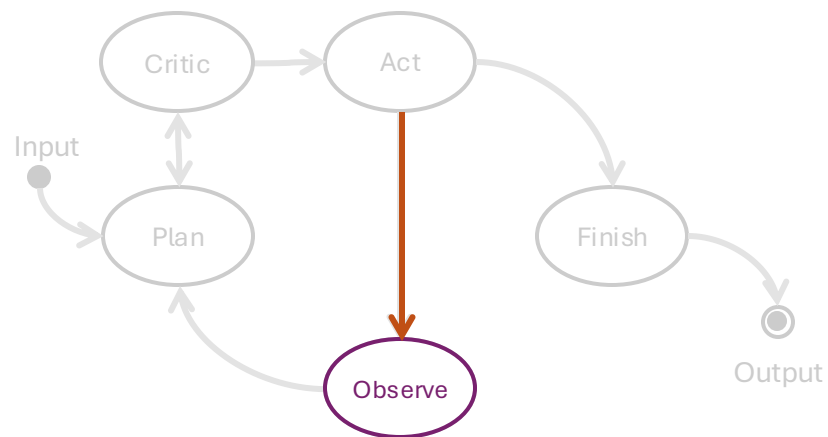
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan example



t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a G

p2:Plan

version: 2
state: Excluded
criticism: "Step for getting a

p3:Plan

version: 3
state: Active
criticism: null
from: s6:Step

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the
state: Finished
action: "Wikipedia[Area of t
action_output: "The United
observation: "UK area is 24
criticism: null

s9:Step

thought: "I need to know the area of the US."
state: Running
action: "Google[Area of the US]"
action_output: "The area of the US is 3,809,525 square miles (9,840,000 square km)"
observation: null

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the
state: Excluded
action: "Wikipedia[Area of t
action_output: "The United
observation: "There is no ar
criticism: "Step doesn't con

s10:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s11:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

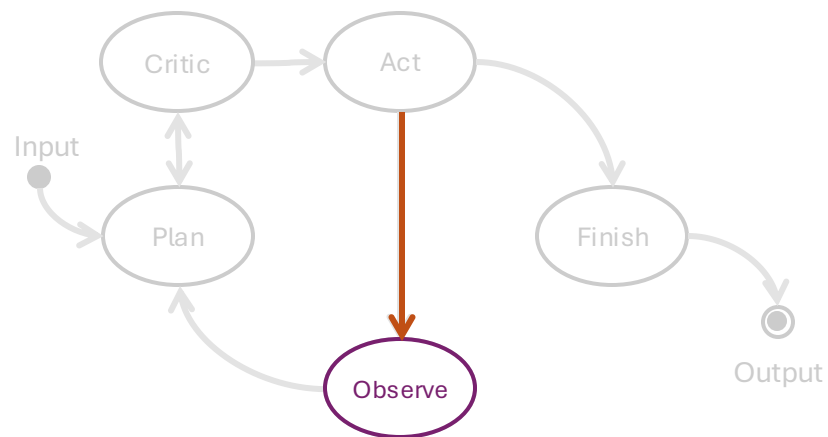
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a G

p2:Plan

version: 2
state: Excluded
criticism: "Step for getting a

p3:Plan

version: 3
state: Active
criticism: null
from: s6:Step

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the
state: Finished
action: "Wikipedia[Area of t
action_output: "The United
observation: "UK area is 24
criticism: null

s9:Step

thought: "I need to know the area of the US."
state: Running
action: "Google[Area of the US]"
action_output: "The area of the US is 3,809,525 square miles (9,840,000 square km)"
observation: "US are is 9,840,000 km2"

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the
state: Excluded
action: "Wikipedia[Area of t
action_output: "The United
observation: "There is no ar
criticism: "Step doesn't cont

s10:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s11:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

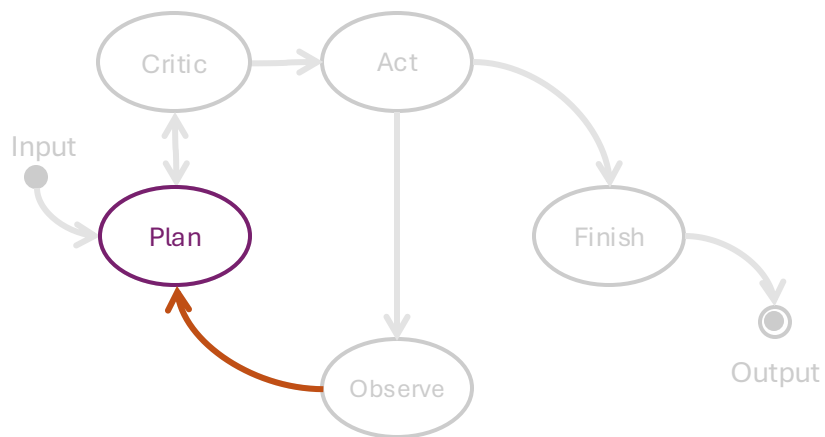
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan example



t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a G

p2:Plan

version: 2
state: Excluded
criticism: "Step for getting a

p3:Plan

version: 3
state: Active
criticism: null
from: s6:Step

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the
state: Finished
action: "Wikipedia[Area of t
action_output: "The United
observation: "UK area is 24
criticism: null

s9:Step

thought: "I need to know the area of the US."
state: Finished
action: "Google[Area of the US]"
action_output: "The area of the US is 3,809,525 square miles (9,840,000 square km)"
observation: "US are is 9,840,000 km2"

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the
state: Excluded
action: "Wikipedia[Area of t
action_output: "The United
observation: "There is no ar
criticism: "Step doesn't con

s10:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s11:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

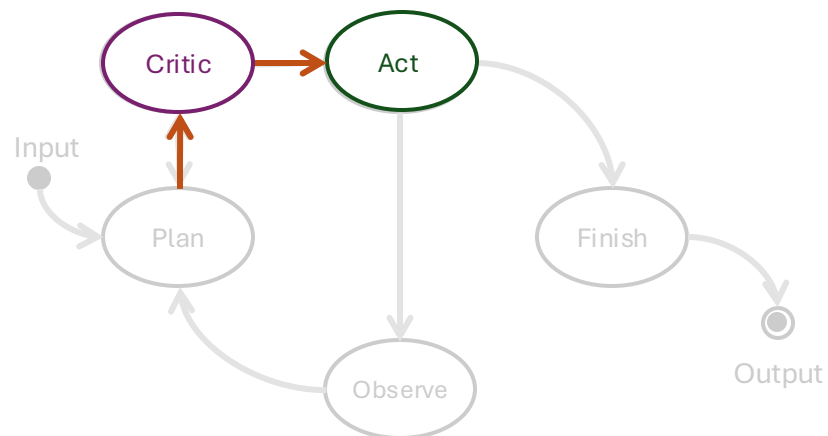
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan example



t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a G

p2:Plan

version: 2
state: Excluded
criticism: "Step for getting a

p3:Plan

version: 3
state: Active
criticism: null
from: s6:Step

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the
state: Finished
action: "Wikipedia[Area of t
action_output: "The United
observation: "UK area is 24
criticism: null

s9:Step

thought: "I need to know the area of the US."
state: Finished
action: "Google[Area of the US]"
action_output: "The area of the US is 3,809,525 square miles (9,840,000 square km)"
observation: "US are is 9,840,000 km2"

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the
state: Excluded
action: "Wikipedia[Area of t
action_output: "The United
observation: "There is no ar
criticism: "Step doesn't cont

s10:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Planned
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s11:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

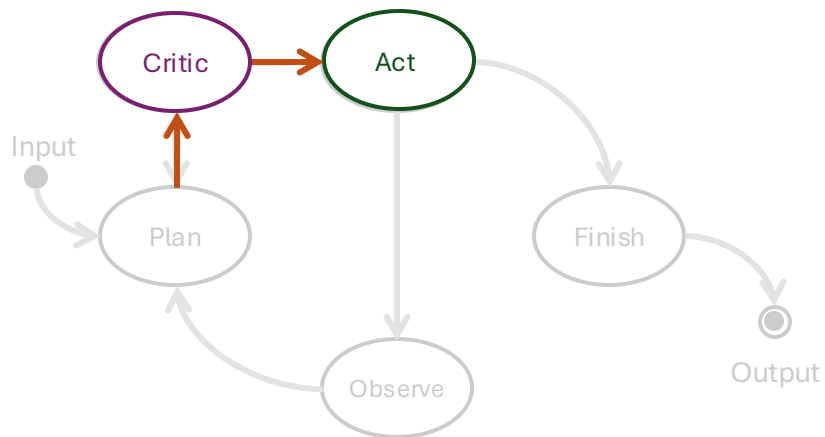
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a G

p2:Plan

version: 2
state: Excluded
criticism: "Step for getting a

p3:Plan

version: 3
state: Active
criticism: null
from: s6:Step

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the
state: Finished
action: "Wikipedia[Area of t
action_output: "The United
observation: "UK area is 24
criticism: null

s9:Step

thought: "I need to know the area of the US."
state: Finished
action: "Google[Area of the US]"
action_output: "The area of the US is 3,809,525 square miles (9,840,000 square km)"
observation: "US are is 9,840,000 km2"

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the
state: Excluded
action: "Wikipedia[Area of t
action_output: "The United
observation: "There is no ar
criticism: "Step doesn't cont

s10:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Running
action: "Python[Calculating the ratio of the areas]"
action_output: null
observation: null

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s11:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

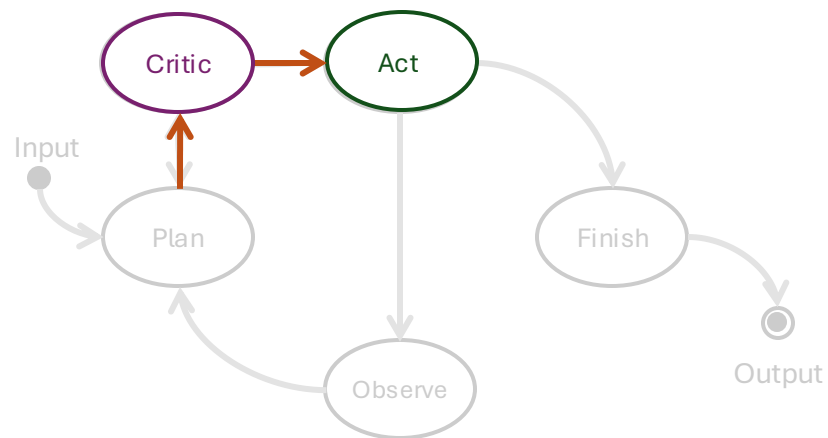
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan example



t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a G

p2:Plan

version: 2
state: Excluded
criticism: "Step for getting a

p3:Plan

version: 3
state: Active
criticism: null
from: s6:Step

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the
state: Finished
action: "Wikipedia[Area of t
action_output: "The United
observation: "UK area is 24
criticism: null

s9:Step

thought: "I need to know the area of the US."
state: Finished
action: "Google[Area of the US]"
action_output: "The area of the US is 3,809,525 square miles (9,840,000 square km)"
observation: "US are is 9,840,000 km2"

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the
state: Excluded
action: "Wikipedia[Area of t
action_output: "The United
observation: "There is no ar
criticism: "Step doesn't con

s10:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Running
action: "Python[Calculating the ratio of the areas]"
action_output: "40.2393033685"
observation: null

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s11:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

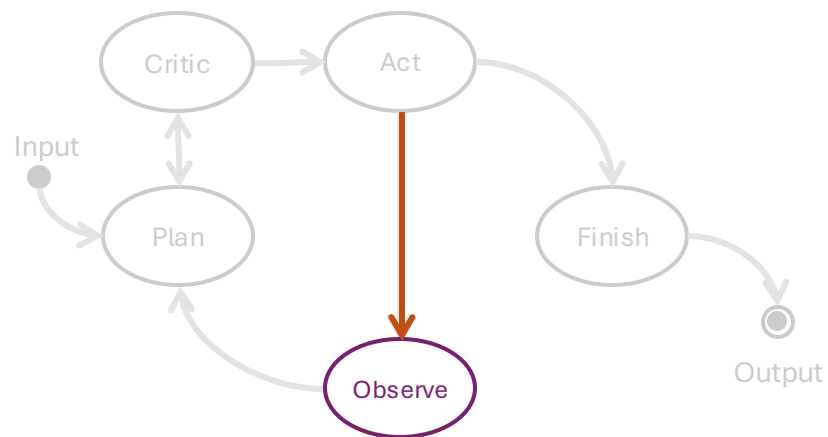
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a G

p2:Plan

version: 2
state: Excluded
criticism: "Step for getting a

p3:Plan

version: 3
state: Active
criticism: null
from: s6:Step

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Area of t
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the
state: Finished
action: "Wikipedia[Area of t
action_output: "The United
observation: "UK area is 24
criticism: null

s9:Step

thought: "I need to know the area of the US."
state: Finished
action: "Google[Area of the US]"
action_output: "The area of the US is 3,809,525 square miles (9,840,000 square km)"
observation: "US are is 9,840,000 km2"

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Area
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the
state: Excluded
action: "Wikipedia[Area of t
action_output: "The United
observation: "There is no ar
criticism: "Step doesn't cont

s10:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Running
action: "Python[Calculating the ratio of the areas]"
action_output: "40.2393033685"
observation: null

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating t
action_output: null
observation: null

s7:Step

thought: "When I have both
state: Planned
action: "Python[Calculating t
action_output: null
observation: null

s11:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

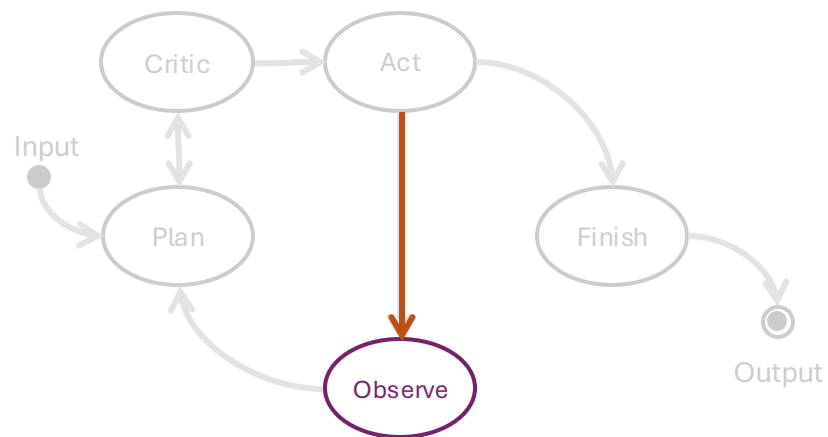
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan example



t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a G

p2:Plan

version: 2
state: Excluded
criticism: "Step for getting a

p3:Plan

version: 3
state: Active
criticism: null
from: s6:Step

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the
state: Finished
action: "Wikipedia[Area of t
action_output: "The United
observation: "UK area is 24
criticism: null

s9:Step

thought: "I need to know the area of the US."
state: Finished
action: "Google[Area of the US]"
action_output: "The area of the US is 3,809,525 square miles (9,840,000 square km)"
observation: "US are is 9,840,000 km2"

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the
state: Excluded
action: "Wikipedia[Area of t
action_output: "The United
observation: "There is no ar
criticism: "Step doesn't con

s10:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Running
action: "Python[Calculating the ratio of the areas]"
action_output: "40.2393033685"
observation: "The result of the ratio is 40.2393033685."

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s11:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

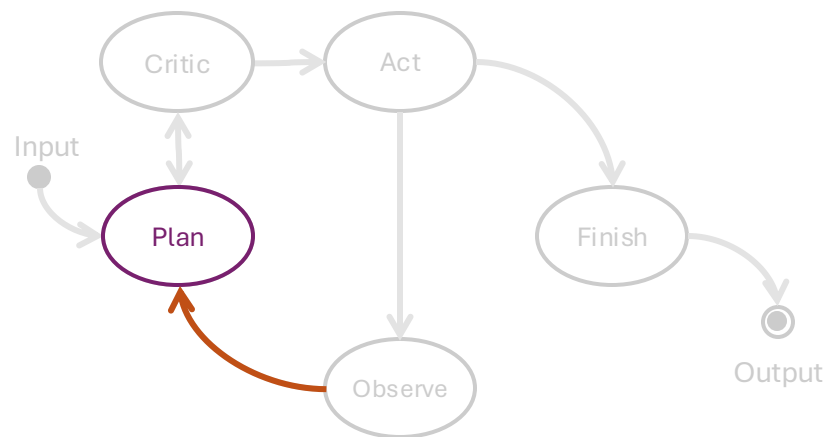
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan example



t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a G

p2:Plan

version: 2
state: Excluded
criticism: "Step for getting a

p3:Plan

version: 3
state: Active
criticism: null
from: s6:Step

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the
state: Finished
action: "Wikipedia[Area of t
action_output: "The United
observation: "UK area is 24
criticism: null

s9:Step

thought: "I need to know the area of the US."
state: Finished
action: "Google[Area of the US]"
action_output: "The area of the US is 3,809,525 square miles (9,840,000 square km)"
observation: "US are is 9,840,000 km2"

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the
state: Excluded
action: "Wikipedia[Area of t
action_output: "The United
observation: "There is no ar
criticism: "Step doesn't con

s10:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Finished
action: "Python[Calculating the ratio of the areas]"
action_output: "40.2393033685"
observation: "The result of the ratio is 40.2393033685."

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s11:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

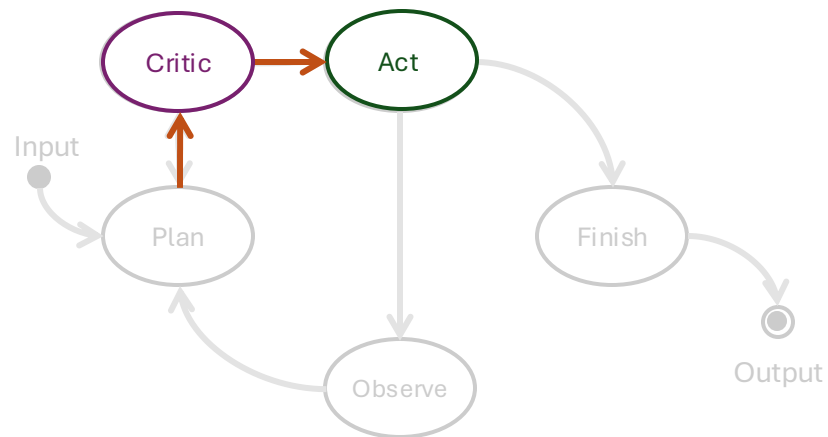
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan example



t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a G

p2:Plan

version: 2
state: Excluded
criticism: "Step for getting a

p3:Plan

version: 3
state: Active
criticism: null
from: s6:Step

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the
state: Finished
action: "Wikipedia[Area of t
action_output: "The United
observation: "UK area is 24
criticism: null

s9:Step

thought: "I need to know the area of the US."
state: Finished
action: "Google[Area of the US]"
action_output: "The area of the US is 3,809,525 square miles (9,840,000 square km)"
observation: "US are is 9,840,000 km2"

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the
state: Excluded
action: "Wikipedia[Area of t
action_output: "The United
observation: "There is no ar
criticism: "Step doesn't con

s10:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Finished
action: "Python[Calculating the ratio of the areas]"
action_output: "40.2393033685"
observation: "The result of the ratio is 40.2393033685."

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s11:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

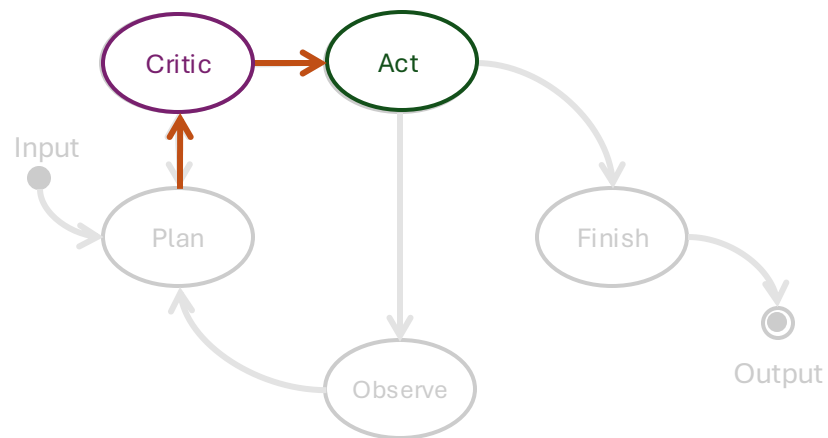
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a G

p2:Plan

version: 2
state: Excluded
criticism: "Step for getting a

p3:Plan

version: 3
state: Active
criticism: null
from: s6:Step

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the
state: Finished
action: "Wikipedia[Area of t
action_output: "The United
observation: "UK area is 24
criticism: null

s9:Step

thought: "I need to know the area of the US."
state: Finished
action: "Google[Area of the US]"
action_output: "The area of the US is 3,809,525 square miles (9,840,000 square km)"
observation: "US are is 9,840,000 km2"

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the
state: Excluded
action: "Wikipedia[Area of t
action_output: "The United
observation: "There is no ar
criticism: "Step doesn't cont

s10:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Finished
action: "Python[Calculating the ratio of the areas]"
action_output: "40.2393033685"
observation: "The result of the ratio is 40.2393033685."

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s11:Step

thought: "Print the final answer."
state: Running
action: "Final[The final number]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

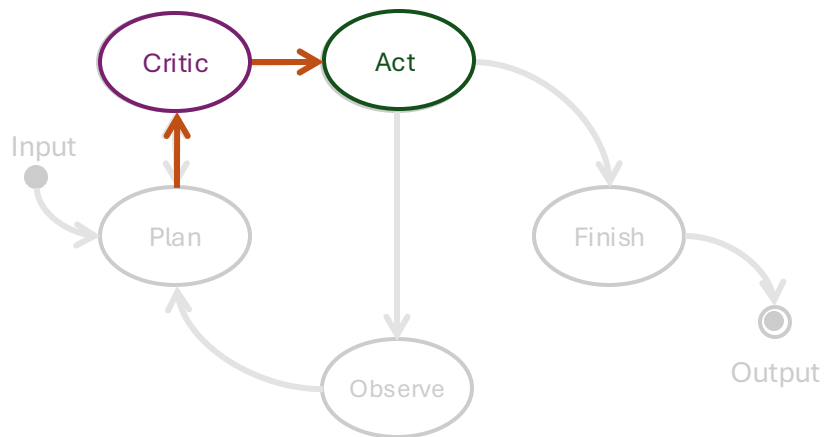
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a G

p2:Plan

version: 2
state: Excluded
criticism: "Step for getting a

p3:Plan

version: 3
state: Active
criticism: null
from: s6:Step

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Area of t
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the
state: Finished
action: "Wikipedia[Area of t
action_output: "The United
observation: "UK area is 24
criticism: null

s9:Step

thought: "I need to know the area of the US."
state: Finished
action: "Google[Area of the US]"
action_output: "The area of the US is 3,809,525 square miles (9,840,000 square km)"
observation: "US are is 9,840,000 km2"

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Area
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the
state: Excluded
action: "Wikipedia[Area of t
action_output: "The United
observation: "There is no ar
criticism: "Step doesn't con

s10:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Finished
action: "Python[Calculating the ratio of the areas]"
action_output: "40.2393033685"
observation: "The result of the ratio is 40.2393033685."

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s11:Step

thought: "Print the final answer."
state: Running
action: "Final[The final number]"
action_output: "Approximately 40.23 times"
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

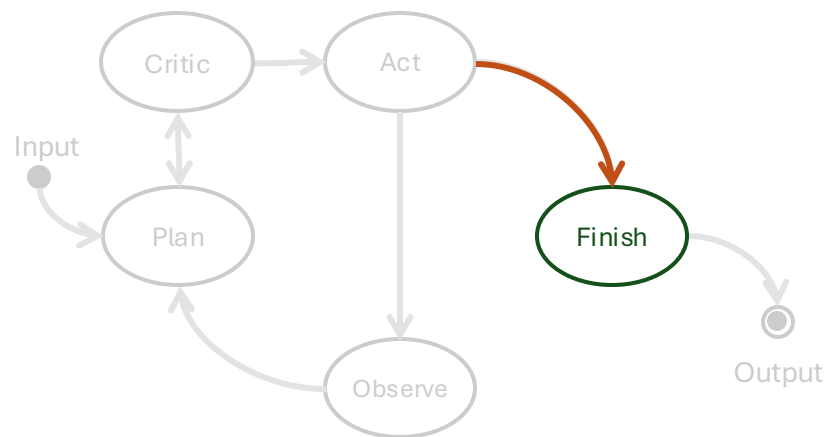
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a G

p2:Plan

version: 2
state: Excluded
criticism: "Step for getting a

p3:Plan

version: 3
state: Active
criticism: null
from: s6:Step

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the
state: Finished
action: "Wikipedia[Area of t
action_output: "The United
observation: "UK area is 24
criticism: null

s9:Step

thought: "I need to know the area of the US."
state: Finished
action: "Google[Area of the US]"
action_output: "The area of the US is 3,809,525 square miles (9,840,000 square km)"
observation: "US are is 9,840,000 km2"

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the
state: Excluded
action: "Wikipedia[Area of t
action_output: "The United
observation: "There is no ar
criticism: "Step doesn't con

s10:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Finished
action: "Python[Calculating the ratio of the areas]"
action_output: "40.2393033685"
observation: "The result of the ratio is 40.2393033685."

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s11:Step

thought: "Print the final answer."
state: Running
action: "Final[The final number]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

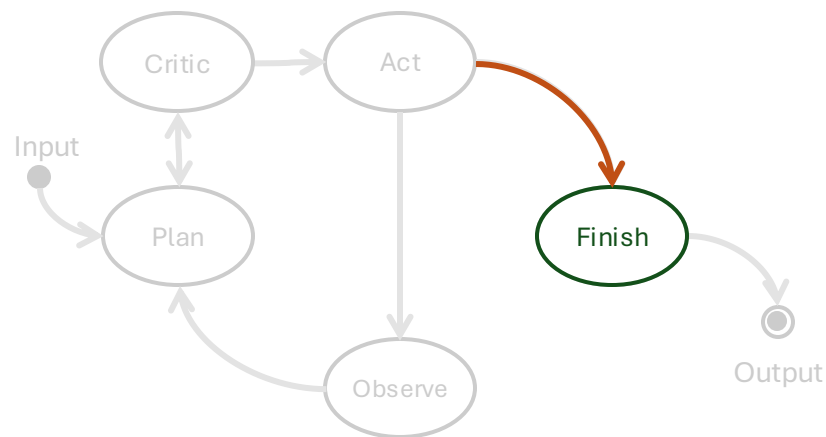
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan example



t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a G

p2:Plan

version: 2
state: Excluded
criticism: "Step for getting a

p3:Plan

version: 3
state: Active
criticism: null
from: s6:Step

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the
state: Finished
action: "Wikipedia[Area of t
action_output: "The United
observation: "UK area is 24
criticism: null

s9:Step

thought: "I need to know the area of the US."
state: Finished
action: "Google[Area of the US]"
action_output: "The area of the US is 3,809,525 square miles (9,840,000 square km)"
observation: "US are is 9,840,000 km2"

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the
state: Excluded
action: "Wikipedia[Area of t
action_output: "The United
observation: "There is no ar
criticism: "Step doesn't cont

s10:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Finished
action: "Python[Calculating the ratio of the areas]"
action_output: "40.2393033685"
observation: "The result of the ratio is 40.2393033685."

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s11:Step

thought: "Print the final answer."
state: Finished
action: "Final[The final number]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

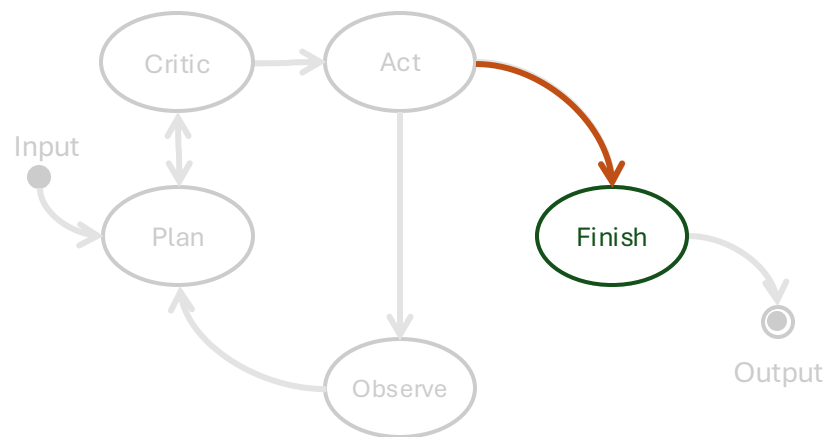
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan example



t1:Task

question: "How many times does the UK fit into the USA?"
state: Uncomplete
answer: null

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a G

p2:Plan

version: 2
state: Excluded
criticism: "Step for getting a

p3:Plan

version: 3
state: Finished
criticism: null
from: s6:Step

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the
state: Finished
action: "Wikipedia[Area of t
action_output: "The United
observation: "UK area is 24
criticism: null

s9:Step

thought: "I need to know the area of the US."
state: Finished
action: "Google[Area of the US]"
action_output: "The area of the US is 3,809,525 square miles (9,840,000 square km)"
observation: "US are is 9,840,000 km2"

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the
state: Excluded
action: "Wikipedia[Area of t
action_output: "The United
observation: "There is no ar
criticism: "Step doesn't con

s10:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Finished
action: "Python[Calculating the ratio of the areas]"
action_output: "40.2393033685"
observation: "The result of the ratio is 40.2393033685."

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s11:Step

thought: "Print the final answer."
state: Finished
action: "Final[The final number]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

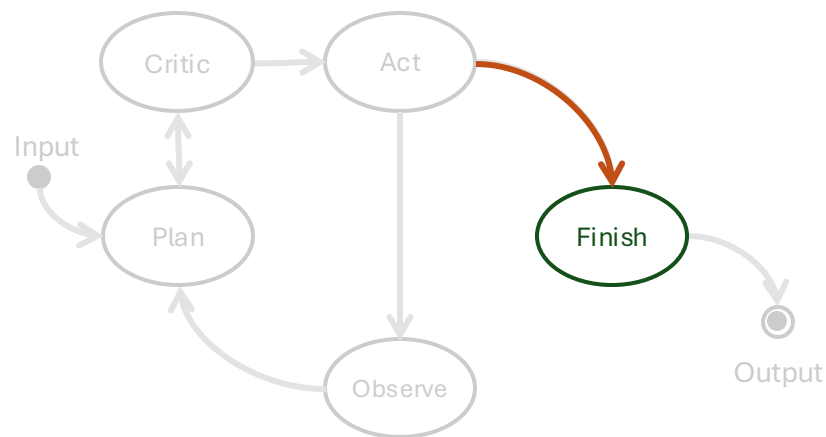
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan example



t1:Task

question: "How many times does the UK fit into the USA?"
state: **Uncomplete**
answer: "Approximately 40.23 times"

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a G

p2:Plan

version: 2
state: Excluded
criticism: "Step for getting a

p3:Plan

version: 3
state: **Finished**
criticism: null
from: s6:Step

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the
state: **Finished**
action: "Wikipedia[Area of t
action_output: "The United
observation: "UK area is 24
criticism: null

s9:Step

thought: "I need to know the area of the US."
state: **Finished**
action: "Google[Area of the US]"
action_output: "The area of the US is 3,809,525 square miles (9,840,000 square km)"
observation: "US are is 9,840,000 km2"

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the
state: Excluded
action: "Wikipedia[Area of t
action_output: "The United
observation: "There is no ar
criticism: "Step doesn't con

s10:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: **Finished**
action: "Python[Calculating the ratio of the areas]"
action_output: "40.2393033685"
observation: "The result of the ratio is 40.2393033685."

s3:Step

thought: "When I have both
state: **Planned**
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both
state: **Planned**
action: "Python[Calculating
action_output: null
observation: null

s11:Step

thought: "Print the final answer."
state: **Finished**
action: "Final[The final number]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: **Planned**
action: "Final[The final num
action_output: null
observation: null

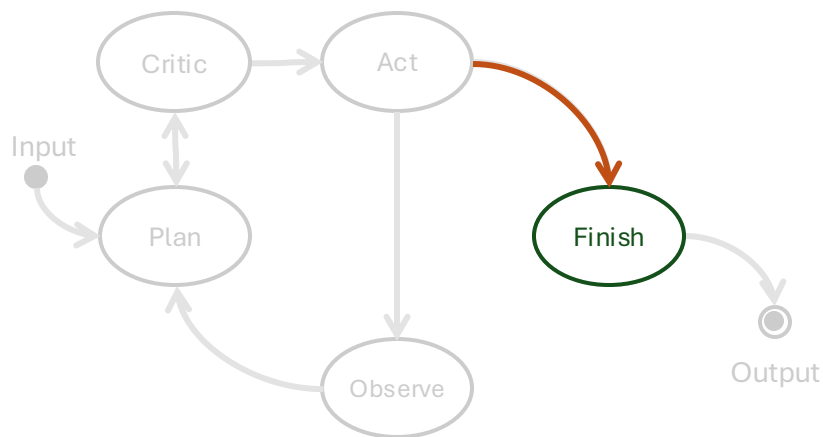
s8:Step

thought: "Print the final answer."
state: **Planned**
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



t1:Task

question: "How many times does the UK fit into the USA?"
state: Completed
answer: "Approximately 40.23 times"

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a G"

p2:Plan

version: 2
state: Excluded
criticism: "Step for getting a"

p3:Plan

version: 3
state: Finished
criticism: null
from: s6:Step

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the
state: Finished
action: "Wikipedia[Area of t
action_output: "The United
observation: "UK area is 24
criticism: null

s9:Step

thought: "I need to know the area of the US."
state: Finished
action: "Google[Area of the US]"
action_output: "The area of the US is 3,809,525 square miles (9,840,000 square km)"
observation: "US are is 9,840,000 km2"

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the
state: Excluded
action: "Wikipedia[Area of t
action_output: "The United
observation: "There is no ar
criticism: "Step doesn't con

s10:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Finished
action: "Python[Calculating the ratio of the areas]"
action_output: "40.2393033685"
observation: "The result of the ratio is 40.2393033685."

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s11:Step

thought: "Print the final answer."
state: Finished
action: "Final[The final number]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

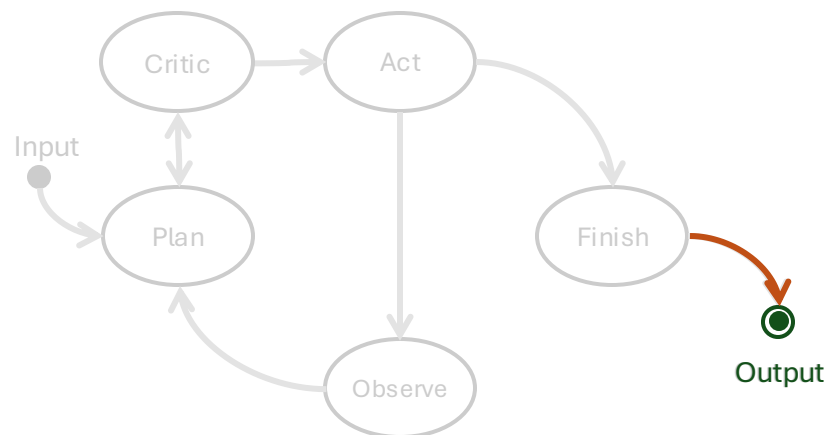
s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null



ReAct agent *evolution*

Critic + Replan *example*



t1:Task

question: "How many times does the UK fit into the USA?"
state: Completed
answer: "Approximately 40.23 times"

p1:Plan

version: 1
state: Excluded
criticism: "You can't use a G

p2:Plan

version: 2
state: Excluded
criticism: "Step for getting a

p3:Plan

version: 3
state: Finished
criticism: null
from: s6:Step

s1:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s5:Step

thought: "I need to know the
state: Finished
action: "Wikipedia[Area of t
action_output: "The United
observation: "UK area is 24
criticism: null

s9:Step

thought: "I need to know the area of the US."
state: Finished
action: "Google[Area of the US]"
action_output: "The area of the US is 3,809,525 square miles (9,840,000 square km)"
observation: "US are is 9,840,000 km2"

s2:Step

thought: "I need to know the
state: Excluded
action: "GeologicService[Are
action_output: null
observation: null
criticism: "GeologicService t

s6:Step

thought: "I need to know the
state: Excluded
action: "Wikipedia[Area of t
action_output: "The United
observation: "There is no ar
criticism: "Step doesn't con

s10:Step

thought: "When I have both areas I can calculate how many times the UK fits into the USA."
state: Finished
action: "Python[Calculating the ratio of the areas]"
action_output: "40.2393033685"
observation: "The result of the ratio is 40.2393033685."

s3:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s7:Step

thought: "When I have both
state: Planned
action: "Python[Calculating
action_output: null
observation: null

s11:Step

thought: "Print the final answer."
state: Finished
action: "Final[The final number]"
action_output: null
observation: null

s4:Step

thought: "Print the final ans
state: Planned
action: "Final[The final num
action_output: null
observation: null

s8:Step

thought: "Print the final answer."
state: Planned
action: "Final[The final number]"
action_output: null
observation: null

The End