



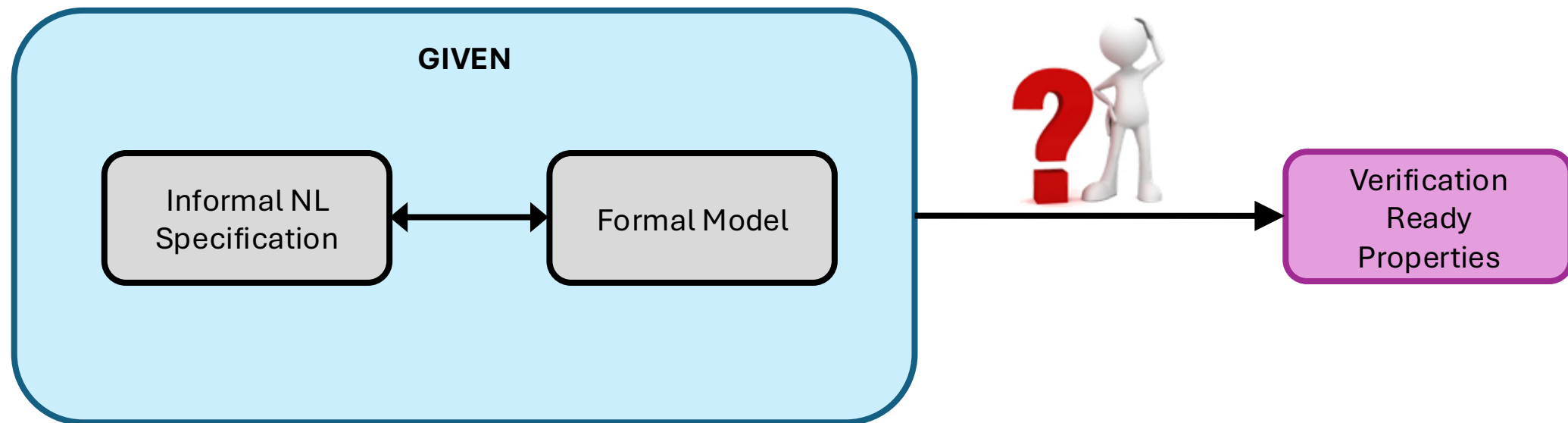
POLITECNICO
MILANO 1863

Towards an Agentic LLM-Based Approach to Requirement Formalization from Unstructured Specifications

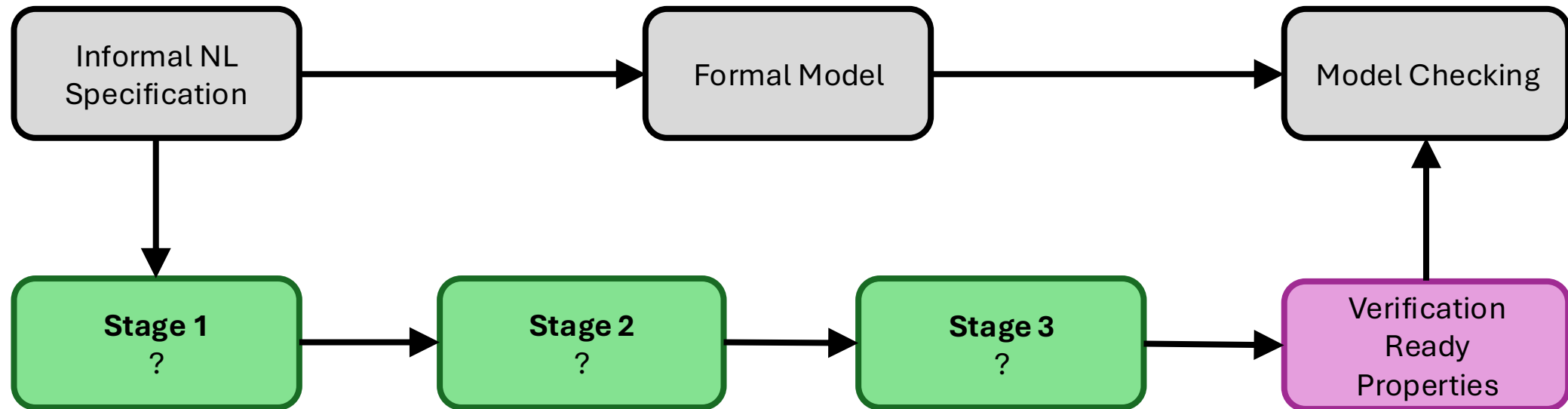
Alberto Tagliaferro, Bruno Guindani, Livia Lestingi, Matteo Rossi

Problem Statement

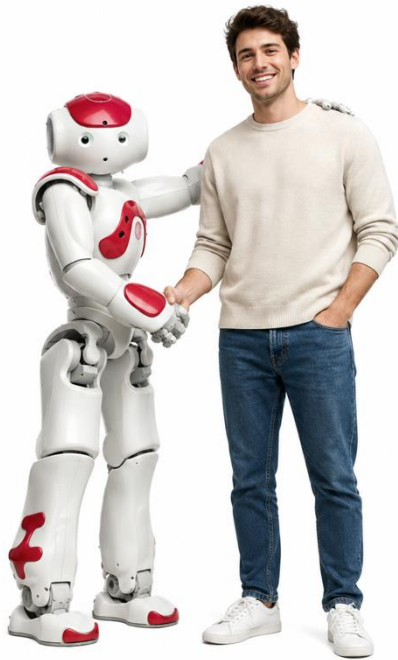
- How can we generate verifiable queries from informal specifications?



Proposed 3-Stage Agentic-Pipeline



Case Study - Introduction



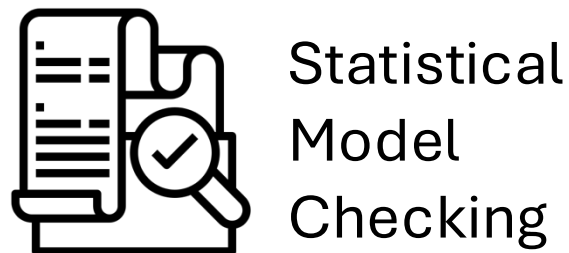
Specification (1/3):

A **human** and a **robot** operate in a **house** with two locations (bedroom, kitchen) and one object initially in the bedroom.

The **task** is strictly sequential: both agents must first **reach the bedroom**. Once co-located there, the robot **picks up** the object. Only after the pick-up is completed does the **human move to the kitchen**, and the **robot follows** while carrying the object.

The **goal** state is reached when both agents are in the kitchen, and the robot still holds the object.

Case Study - Introduction



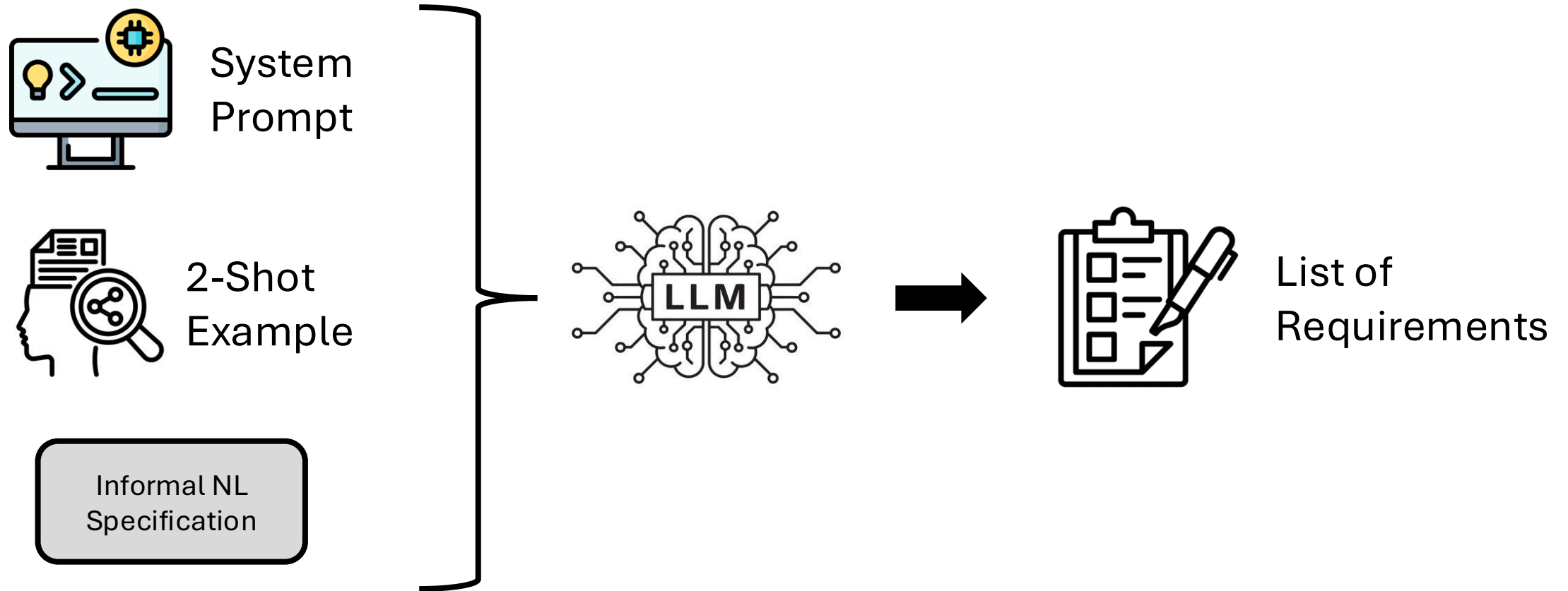
Specification (1/3):

A **human** and a **robot** operate in a **house** with two locations (bedroom, kitchen) and one object initially in the bedroom.

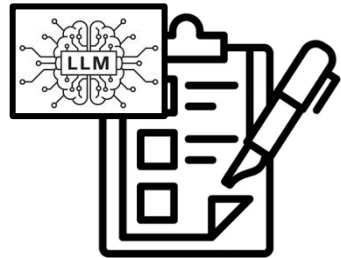
The **task** is strictly sequential: both agents must first **reach the bedroom**. Once co-located there, the robot **picks up** the object. Only after the pick-up is completed does the **human move to the kitchen**, and the **robot follows** while carrying the object.

The **goal** state is reached when both agents are in the kitchen, and the robot still holds the object.

Stage 1 – Requirement Extraction



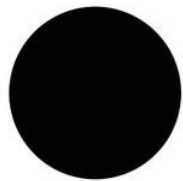
Stage 1 – Requirement Extraction



Generated
List of
Requirements



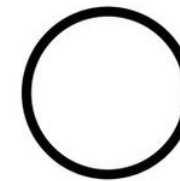
Ground-Truth
List of
Requirements



Perfect Match (43.6%)



Partial Match (38.2%)

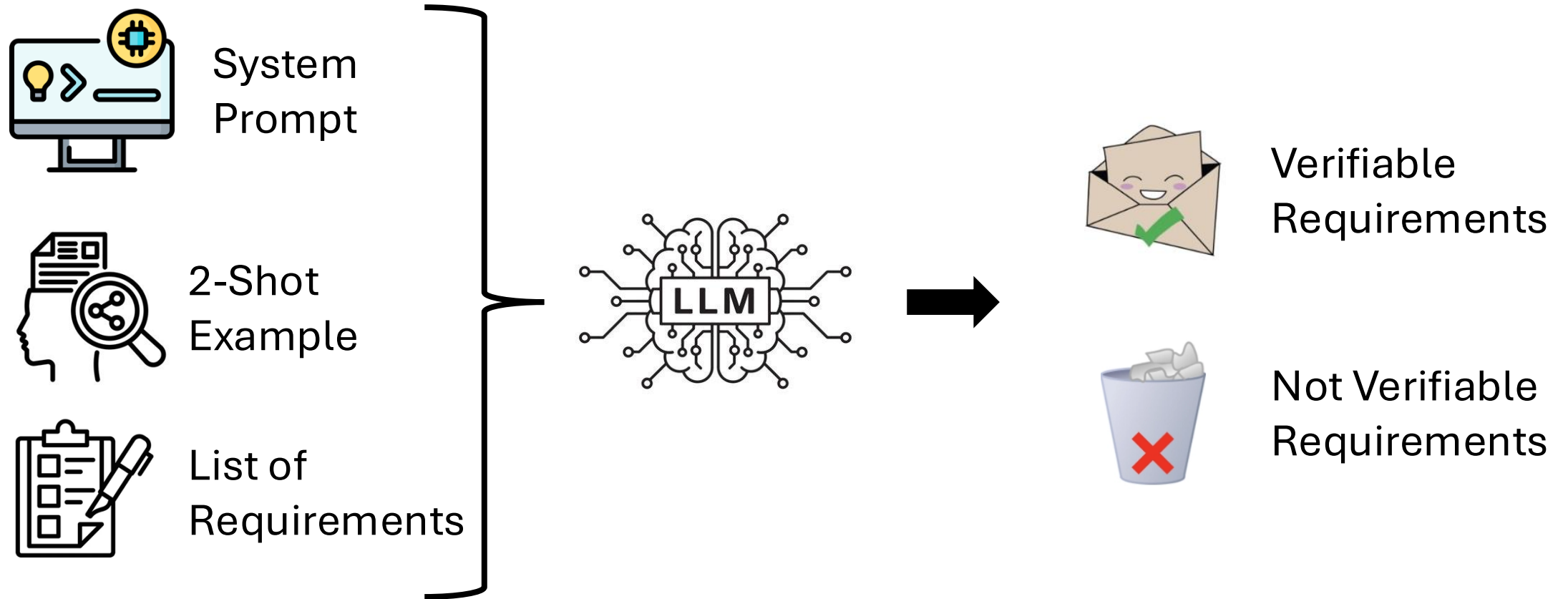


No Match (18.2%)

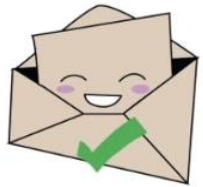
- The Robot must reach the bedroom
- The Human must reach the bedroom

- All the agents must reach the bedroom

Stage 2 – Requirement Classification



Stage 2 – Requirement Classification



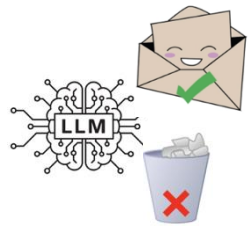
Verifiable
Requirements



Not Verifiable
Requirements

- Model Capabilities:
 - The item cannot fall during transportation
- UPPAAL SMC Limitations:
 - The human shall remain in the bedroom until the robot has successfully grasped the object
 - The robot must reach the coffee machine before preparing the coffee.

Stage 2 – Requirement Classification



Generated
Classification

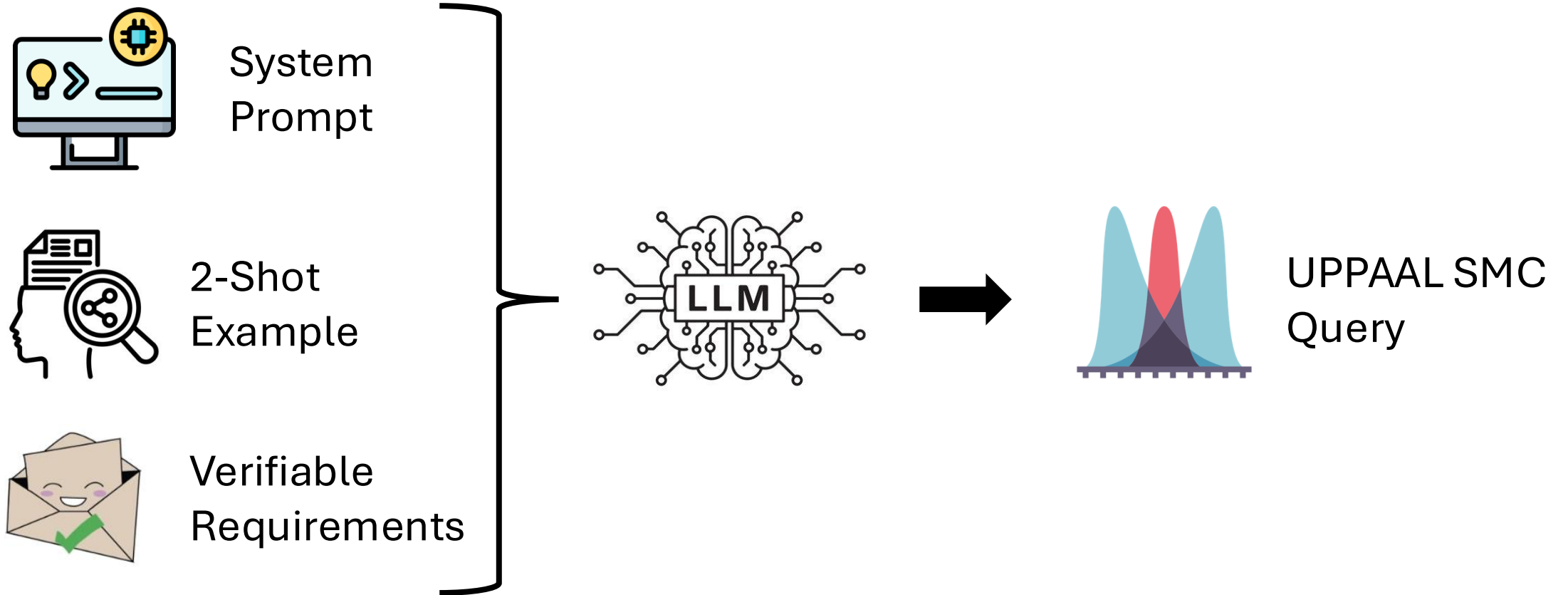


Ground-Truth
Classification

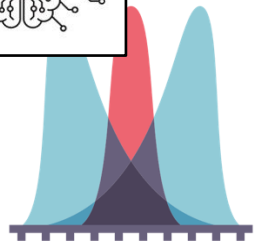
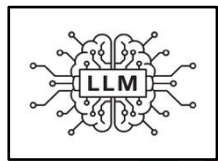
	Predicted Verifiable	Predicted Unverifiable
True Verifiable	TP = 49	FN = 3
False Verifiable	FP = 5	TN = 14

Accuracy	88.7%
Precision	90.7%
Recall	94.2%

Stage 3 – Formal Translation



Stage 3 – Formal Translation



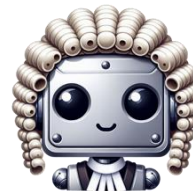
Generated
UPPAAL SMC
Query

1. Syntax Check (95.8%)

2. Ground-Truth Comparison:

1. Exact Match (34.7%)

2. Semantic Match (70.5%)



LLM as Judge

Total Accuracy
77.8%

Summary

