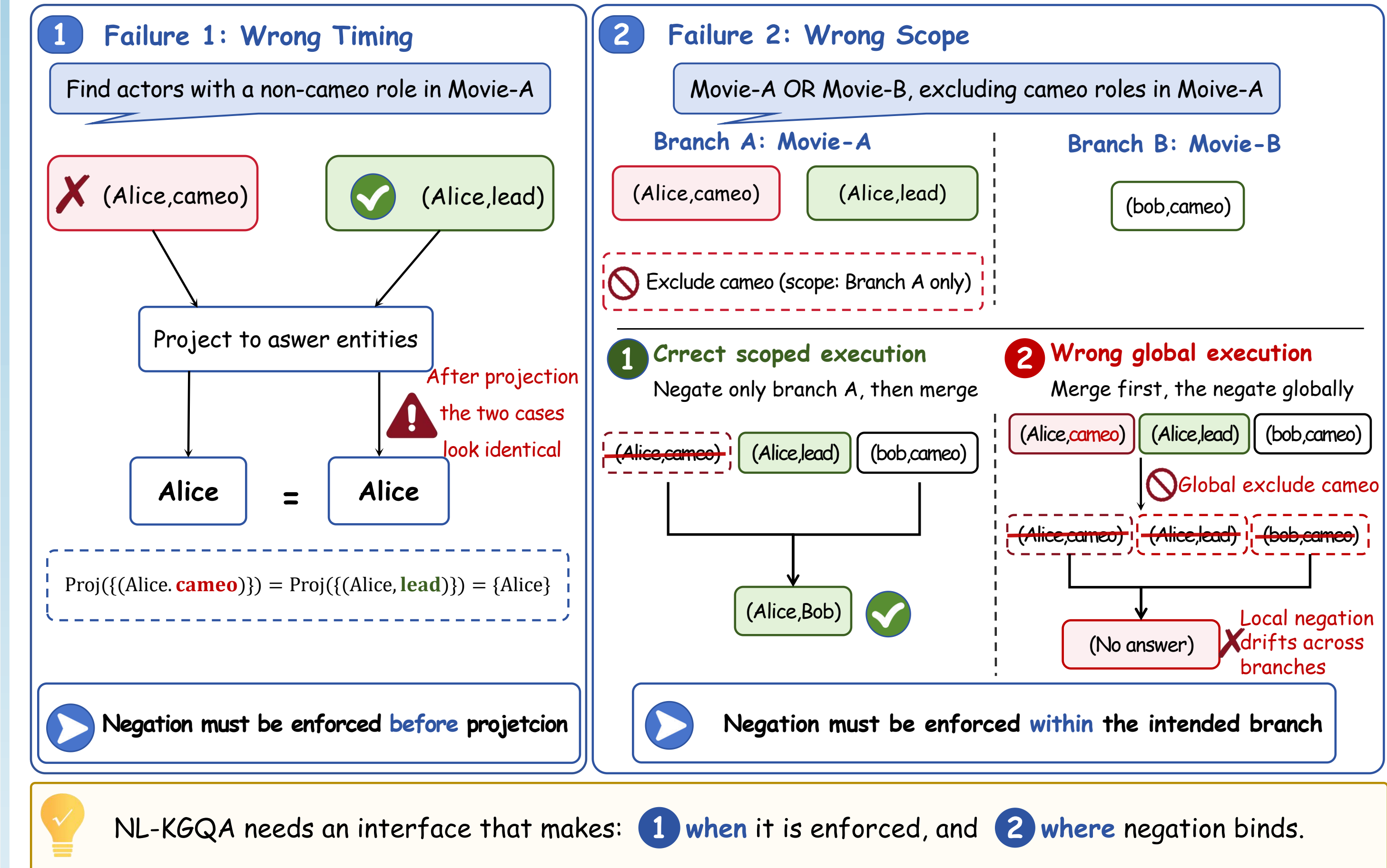


Motivation and Contribution

Motivation: To make natural-language negation in KGQA, preventing errors caused by incorrect negation scope or timing.



Key Contributions: We introduce the Executable Exchange Contract (EEC) and instantiate it for CQA and NL-KGQA.

- **Executable Exchange Contract.** We formulate negation in LLM-to-KG reasoning as an exchange contract between a specifier and an executor, separating what to exclude (with scope) from how it is executed.
- **Scoped Enforcement Boundary.** We identify the $P \rightarrow N \rightarrow C/D$ enforcement boundary, showing that the exclusion scope must be enforced before projection and branch merging.
- **LM-MATLOGIC Instantiation.** We implement MATLOGIC, a matrix-based executor that compiles scoped exclusions into scope-local masks and returns witness-backed traces for verification.

Definition and Formulation

Given a natural-language question q and a knowledge graph G , the specifier constructs an executable EEC request:

$$EEC_{req} = \langle T, E_q, N \rangle,$$

where T is an operator skeleton, E_q denotes linked anchor entities, and N encodes the exclusion constraints. Each exclusion constraint is defined as

$$d = \langle \text{scope_id}, \phi_d \rangle \in D,$$

where scope_id selects the intermediate state under exclusion, and ϕ_d is the executable predicate over its evidence space. The corresponding response:

$$EEC_{resp} = \langle A_{ans}, \tau \rangle,$$

where A_{ans} denotes the final answers and τ is a witness-backed trace that supports verification of the returned answers.

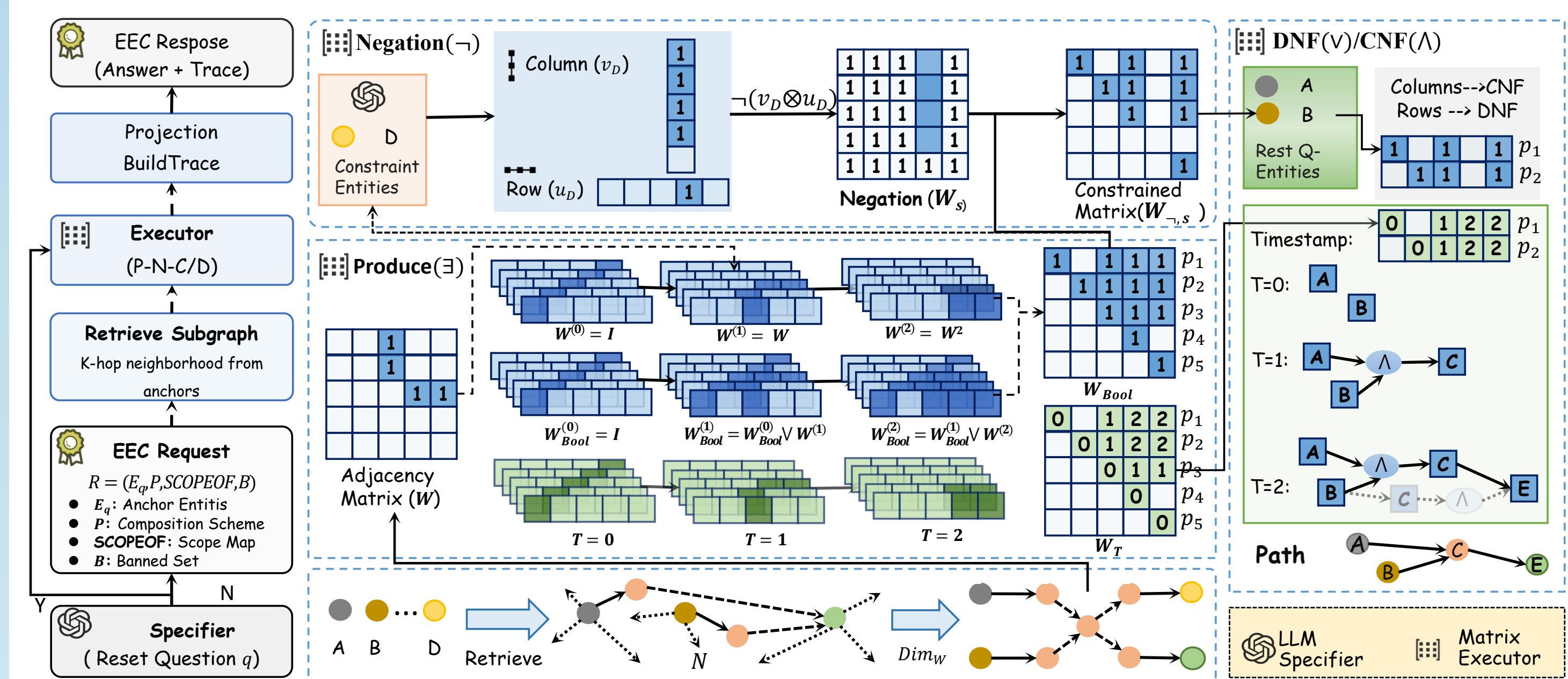
EEC & LLM-MATLOGIC

1. EEC INSTANTIATION. EEC turns natural-language exclusions into explicit, scoped, and executable.

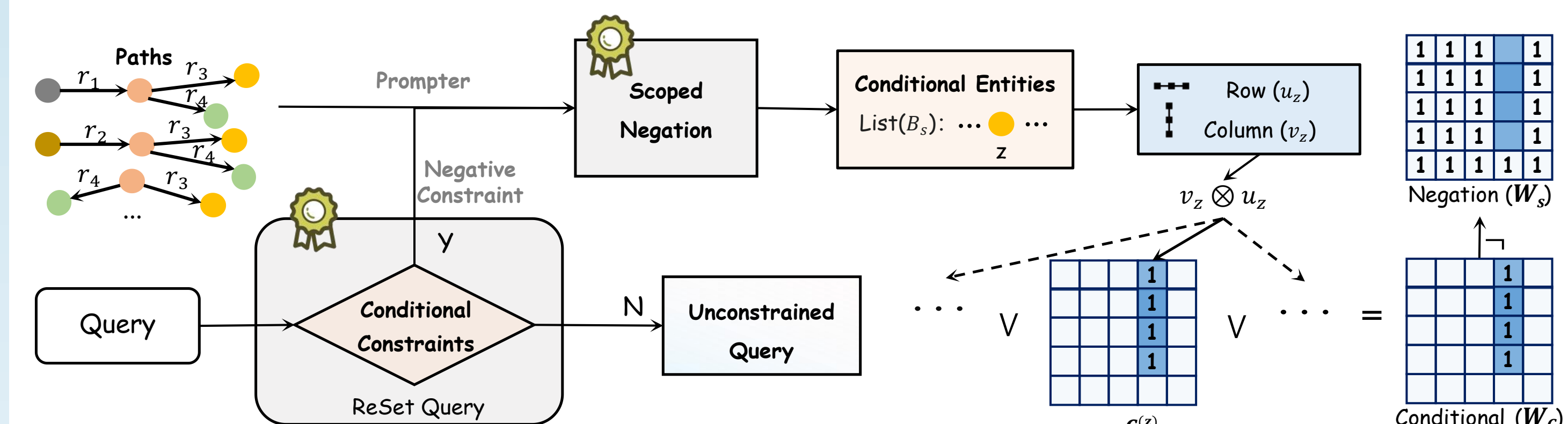
$$R = \langle E_q, P, \text{SCOPEOF}, B \rangle,$$

where E_q are anchors, P is the composition, SCOPEOF set scopes, and B gives scoped banned sets.

2. LLM-MATLOGIC FRAMEWORK. LLM-MATLOGIC separates LLM specification from MATLOGIC execution.



3. SCOPE-LOCAL NEGATION. Enforces localized exclusions.

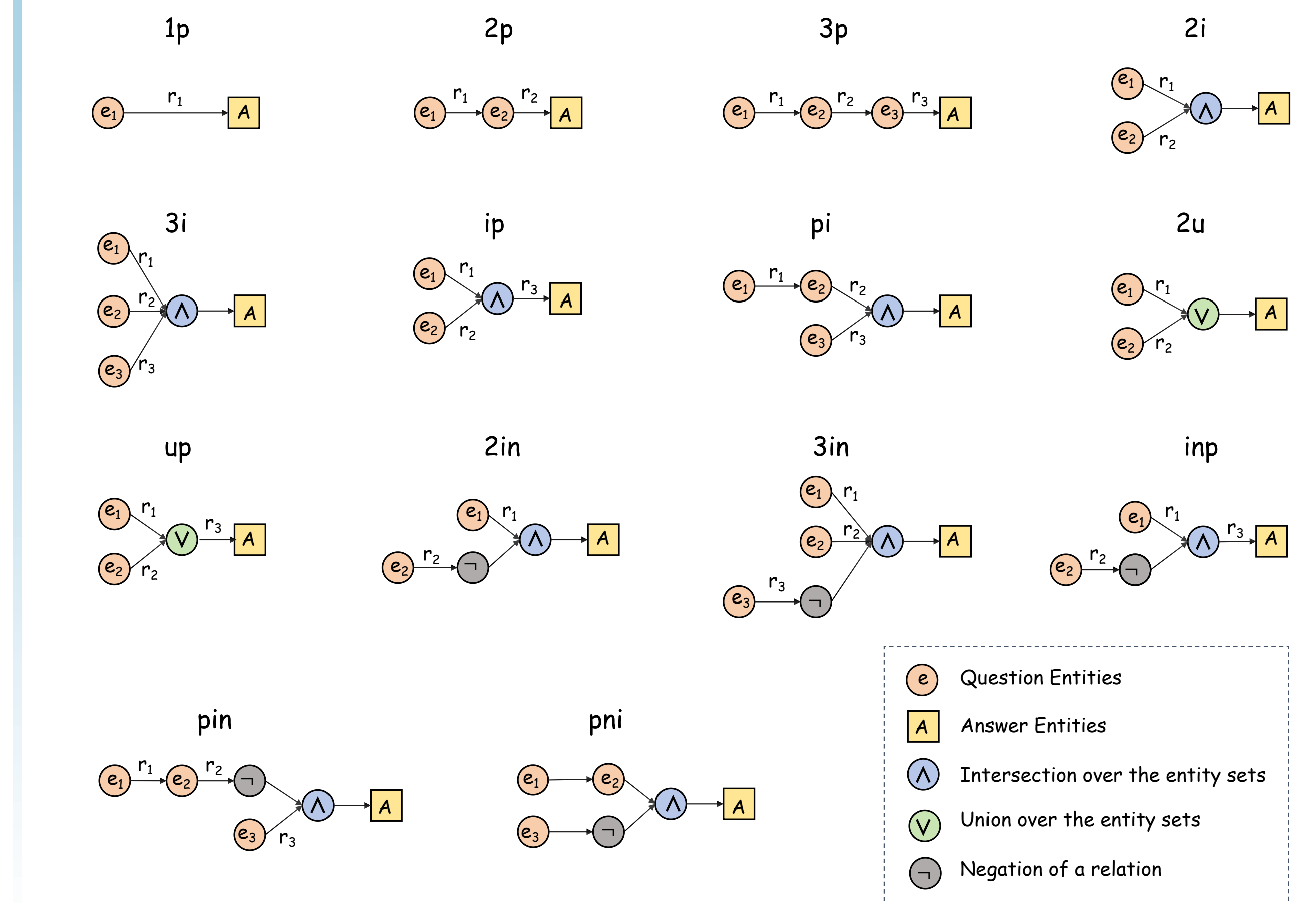


4. UNIFIED PARADIGM. A enforcement paradigm $P \rightarrow N \rightarrow C/D$.

Orders	Applicability / limitation
$P \rightarrow N \rightarrow C/D$	<i>Always valid</i> for scoped exclusions: Global exclusions can be localized to scope negation before composition.
$P \rightarrow C/D \rightarrow N$	Valid for <i>negation-free</i> queries; with negation, post-composition may cause witness loss or scope drift.
$N \rightarrow P \rightarrow C/D$	<i>Not meaningful</i> : for multi-hop structures, N has no target scope or scoped witness evidence to enforce before P .
$N \rightarrow C/D \rightarrow P$	<i>Ill-posed</i> : C/D consumes witness evidence produced by P .
$C/D \rightarrow N \rightarrow P$	<i>Ill-posed</i> for the same reason.
$C/D \rightarrow P \rightarrow N$	<i>Ill-posed</i> ; also corresponds to post-hoc filtering after witness consumption.

NOTE: Ill-posed means the required intermediate evidence is not yet available.

Fourteen Query structure



Experiment 1 and Results

RQ1. Can LLM-MATLOGIC handle diverse query structures?

Type	Model	1p	2p	3p	2i	3i	ip	pi	2u	up
FB15k-237										
Embedding	BetaE	38.7	12.5	10.3	27.5	42.7	23.1	14.1	12.9	10.6
Prompting	LARK	65.6	34.7	21.8	39.7	39.3	21.4	42.8	47.4	23.5
Prompting	MFC	67.9	35.2	30.9	42.4	43.6	27.6	37.7	48.2	20.1
EEC	LLM-MATLOGIC	69.2	58.2	42.5	57.1	43.2	31.8	24.2	39.6	26.3
NELL995										
Embedding	BetaE	54.1	13.8	10.9	38.5	46.7	25.2	14.4	13.1	8.9
Prompting	LARK	73.2	38.3	25.5	46.1	42.2	23.5	26.8	26.9	36.7
Prompting	MFC	75.4	39.6	32.9	49.5	46.8	28.6	28.3	33.1	31.6
EEC	LLM-MATLOGIC	79.4	64.5	44.2	59.7	44.5	31.9	22.7	24.4	35.2

Model	2in	3in	inp	pin	pni
FB15k-237					
BetaE	5.9	8.2	7.9	4.1	2.8
LARK	7.5	4.8	23.3	18.7	3.6
LLM-MATLOGIC	11.4	12.7	15.1	7.5	10.9
NELL995					
BetaE	5.7	7.5	9.6	3.3	3.1
LARK	10.6	6.9	24.4	14.5	7.8
LLM-MATLOGIC	13.8	15.2	17.9	6.7	13.3

Experiment 2 and Results

RQ2. Can EEC be applied to real-world NL-KGQA??

Model	CWQ		WebQSP	
	Hits@1	Calls	Hits@1	Calls
IO-prompting	45.2	1	67.2	1
CoT-prompting	47.0	1	67.7	1
ToG	51.8	14.3	72.2	10.7
MFC	62.8	9.7	78.9	7.5
LLM-MATLOGIC	63.5	3.0	76.7	3.0

LLM-MATLOGIC is evaluated on *real NL-KGQA tasks*: the LLM converts questions into executable EEC requests, while MATLOGIC performs unified structural logic reasoning over retrieved KG subgraphs.