

(Description) Logics for Information Modelling and Access

Enrico Franconi

`franconi@cs.man.ac.uk`

`http://www.cs.man.ac.uk/~franconi`

Department of Computer Science, University of Manchester

Univ. of Liverpool – 22 January 2002

Summary

- Description Logics
- The role of logics in Information Systems
- Conceptual Modelling and Query Management

Description Logics – the standard view

- Expressive decidable fragments of (first-order) classical logic
- Close correspondence with modal logics (e.g., \mathcal{ALC} vs. \mathbf{K})
- Sound and complete algorithms implemented in efficient reasoners
- Knowledge representation formalism derived by semantic networks and frames in Artificial Intelligence
- Close correspondence with well known database conceptual data models

Knowledge representation is about objects

Description logics describe classes of objects (concepts) and their inter-relationships (roles).

The \mathcal{ALC} concept expression

$\text{Professor} \sqcap \exists \text{TEACHES}. \text{UG-Course} \sqcap \forall \text{TEACHES}. \text{CS-Course}$

corresponds to the \mathbf{K} formula

$\text{Professor} \wedge \Diamond \text{UG-Course} \wedge \Box \text{CS-Course}$

where the accessibility relation is interpreted as the *TEACHES* relation

Description Logics are multi-modal

The \mathcal{ALC} concept expression

$\text{Professor} \sqcap \exists \text{TEACHES}. \text{UG-Course} \sqcap \exists \text{DEGREE}. Bs$

corresponds to the \mathbf{K}_m formula (over the same object domain)

$\text{Professor} \wedge \Diamond_{\text{TEACHES}} \text{UG-Course} \wedge \Diamond_{\text{DEGREE}} Bs$

Modalities (as roles) may have different properties

The \mathcal{ALC} concept expression

$\text{Professor} \sqcap \exists \text{TEACHES}. \text{UG-Course} \sqcap$
 $\exists \text{IS-PART}. (\text{Staff} \sqcap \exists \text{IS-LOCATED}. \text{Department})$

corresponds to the $\mathbf{K}_m \cup \mathbf{K4}_m$ formula (over the same object domain)

$\text{Professor} \wedge \Diamond_{\text{TEACHES}} \text{UG-Course} \wedge$
 $\Diamond_{\text{IS-PART}} (\text{Staff} \wedge \Diamond_{\text{IS-LOCATED}} \text{Department})$

where TEACHES is a \mathbf{K}_m modality and IS-PART , IS-LOCATED are $\mathbf{K4}_m$ modalities

Relational structures

- “Modal logics are not appropriate as a representational tool since they do not always capture the details of the models”
- Do we care?

Relational structures

- “Modal logics are not appropriate as a representational tool since they do not always capture the details of the models”
- Do we care?
- No: this is not a bug, it is a feature!

Why?

Relational structures

- “Modal logics are not appropriate as a representational tool since they do not always capture the details of the models”
- Do we care?
- No: this is not a bug, it is a feature!
Why?
- If a formula is satisfiable in a model, it is also satisfiable in a model with the indistinguishable property. Hence, reasoning is not affected.

Additional genuine modalities

- Time, space, belief, etc: combination of modal logics over *distinct* domains (i.e., the object and the modal domains)
- Example of $\mathcal{ALC} \cup \text{LTL}$ concept expression:
 $\text{Professor} \sqcap \Diamond(\exists \text{TEACHES. UG-Course}) \sqcap \forall \Box \text{TEACHES. CS-Course}$
- Asymmetric extension

Global axioms

$$\exists \text{TEACHES. Course} \sqsubseteq (\text{Student} \sqcap \exists \text{DEGREE. Bs}) \sqcup \text{Prof}$$

$$\text{Prof} \sqsubseteq \exists \text{DEGREE. Ms}$$

$$\exists \text{DEGREE. Ms} \sqsubseteq \exists \text{DEGREE. Bs}$$

$$\text{Ms} \sqcap \text{Bs} \sqsubseteq \perp$$

- Axioms should be satisfied by each object in the domain
- Satisfiability and logical implication in $\mathcal{ALC}(\mathbf{K}_m)$ become EXPTIME-complete

Global axioms, II

- $\mathbf{K}_m^{\mathcal{H}}$ extends \mathbf{K}_m with statements on inclusions between modalities
- Decision problems for $\mathbf{K}_m^{\mathcal{H}}$ and $\mathbf{K4}_m$ are in PSPACE
- The *universal modality* can be encoded in $\mathbf{K}_m^{\mathcal{H}} \cup \mathbf{K4}_m$, and axioms can be internalised:
 - Define new *transitive* modality U that *includes* all other modalities
 - Satisfiability of ϕ w.r.t. $\psi_1 \rightarrow \varphi_1, \dots, \psi_n \rightarrow \varphi_n$ is equivalent to satisfiability of
$$\phi \wedge \Box_U((\psi_1 \rightarrow \varphi_1) \wedge \dots \wedge (\psi_n \rightarrow \varphi_n))$$
- Satisfiability and logical implication in $\mathbf{K}_m^{\mathcal{H}} \cup \mathbf{K4}_m$ are EXPTIME-complete
- FaCT implements $\mathbf{K}_m^{\mathcal{H}} \cup \mathbf{K4}_m$

n -ary Relations

- Relations between objects in the world may necessarily involve more than just two objects
- Full fledged relational structures are needed, beyond Kripke structures
- We want to maintain the modal logic flavour
- \mathcal{DLR} properly extends \mathcal{ALC} with n -ary relations

\mathcal{DLR}

$$R \rightarrow \top_n \mid RN \mid \neg R \mid R_1 \sqcap R_2 \mid R_1 \sqcup R_2 \mid U_i/n : C$$

$$C \rightarrow \top \mid CN \mid \neg C \mid C_1 \sqcap C_2 \mid C_1 \sqcup C_2 \mid \exists[U_i]R \mid \exists^{\leq k}[U_i]R$$

$\text{Works-for} \sqsubseteq \text{subj}/2 : \text{Employee} \sqcap \text{obj}/2 : \text{Project}$

$\text{Manager} \sqsubseteq \text{Employee} \sqcap \neg \exists[\text{subj}]\text{Works-for}$

\mathcal{DLR} includes \mathcal{ALCQI} : if R is a binary relation (i.e., a *role*) with named attributes *first* and *second* then

$$\exists R. C \equiv \exists[\textit{first}](R \sqcap (\textit{second}/2 : C))$$

Reasoning in \mathcal{DLR} is EXPTIME-complete

\mathcal{DLR} syntax	\mathcal{DLR} semantics	RD encoding
\top_n	$\top_n^{\mathcal{I}} \subseteq (\Delta^{\mathcal{I}})^n$	\top_n
RN	$RN^{\mathcal{I}} \subseteq \top_n^{\mathcal{I}}$	C_{RN}
$\neg R$	$\top_n^{\mathcal{I}} \setminus R^{\mathcal{I}}$	$\neg C_R$
$R_1 \sqcap R_2$	$R_1^{\mathcal{I}} \cap R_2^{\mathcal{I}}$	$C_{R_1} \sqcap C_{R_2}$
$U_i/n : C$	$\{\langle d_1, \dots, d_n \rangle \in \top_n^{\mathcal{I}} \mid d_i \in C^{\mathcal{I}}\}$	$\top_n \sqcap \forall U_i. C$
\top	$\top^{\mathcal{I}} = \Delta^{\mathcal{I}}$	\top
CN	$CN^{\mathcal{I}} \subseteq \top^{\mathcal{I}}$	CN
$\neg C$	$\top^{\mathcal{I}} \setminus C^{\mathcal{I}}$	$\neg C$
$C_1 \sqcap C_2$	$C_1^{\mathcal{I}} \cap C_2^{\mathcal{I}}$	$C_1 \sqcap C_2$
$\exists[U_i]R$	$\{d \in \top^{\mathcal{I}} \mid \exists \langle d_1, \dots, d_n \rangle \in R. d_i = d\}$	$\exists U_i^-. C_R$
$\exists^{\leq k}[U_i]R$	$\{d \in \top^{\mathcal{I}} \mid \#\{\langle d_1, \dots, d_n \rangle \in R \mid d_i = d\} \leq k\}$	$\leq k U_i^-. C_R$

Encoding conceptual data models in \mathcal{DLR}

- Object-oriented data models (e.g., UML and ODMG)
- Semantic data models (e.g., EER and ORM)
- Frame-based ontology languages (e.g., OIL)

Encoding conceptual data models in \mathcal{DLR}

- Object-oriented data models (e.g., UML and ODMG)
- Semantic data models (e.g., EER and ORM)
- Frame-based ontology languages (e.g., OIL)
- Theorems **prove** that a conceptual schema and its encoding as \mathcal{DLR} inclusion dependencies constrain every database state in the same way – i.e., the models of the \mathcal{DLR} theory correspond to the legal database states of the conceptual schema, and vice-versa.

Classical Integrity Constraints in \mathcal{DLR}

- arbitrary *boolean constructs*
- *unary inclusion* dependencies (e.g., referential integrity)
- special forms of *typed inclusion* dependencies
- *existence* and *exclusion* dependencies
- *unary functional* dependencies
- *view definitions*

Extensions of \mathcal{DLR}

\mathcal{DLR}_{reg} : regular expressions and recursive views (beyond FOL)

\mathcal{DLR}_{US} : combination with temporal constructs to model temporal databases

\mathcal{DLR}_{key} : general key constraints

Queries under \mathcal{DLR} constraints

- A query is an open FOL formula, whose predicates may be constrained by a \mathcal{DLR} theory
- We consider only the conjunctive existential fragment (the conjunctive queries, or non-recursive datalog queries)

- Example:

$$Q_1(x, y) \text{ :- } (\neg \text{Professor})(x) \wedge \text{TEACHES}(x, y) \wedge \\ (\text{UG-Course} \sqcup \text{CS-Course})(x)$$

Semantics of Evaluation and Containment

The *evaluation* of a query Q of arity n given a \mathcal{DLR} theory Σ over a model \mathcal{I} satisfying Σ

Semantics of Evaluation and Containment

The *evaluation* of a query Q of arity n given a \mathcal{DLR} theory Σ over a model \mathcal{I} satisfying Σ

is the set $\text{ANS}(Q, \mathcal{I})$ composed by n -tuples \vec{o} such that

Semantics of Evaluation and Containment

The *evaluation* of a query Q of arity n given a \mathcal{DLR} theory Σ over a model \mathcal{I} satisfying Σ

is the set $\text{ANS}(Q, \mathcal{I})$ composed by n -tuples \vec{o} such that

$$\text{ANS}(Q, \mathcal{I}) = \{ \vec{o} \mid \mathcal{I} \models \bigvee_j \exists \vec{y}_j. Q_j(\vec{o}, \vec{y}_j, \vec{c}_j) \}$$

Semantics of Evaluation and Containment

The *evaluation* of a query Q of arity n given a \mathcal{DLR} theory Σ over a model \mathcal{I} satisfying Σ

is the set $\text{ANS}(Q, \mathcal{I})$ composed by n -tuples \vec{o} such that

$$\text{ANS}(Q, \mathcal{I}) = \{ \vec{o} \mid \mathcal{I} \models \bigvee_j \exists \vec{y}_j. Q_j(\vec{o}, \vec{y}_j, \vec{c}_j) \}$$

Q_1 *is contained* in Q_2 with respect to a \mathcal{DLR} theory Σ

Semantics of Evaluation and Containment

The *evaluation* of a query Q of arity n given a \mathcal{DLR} theory Σ over a model \mathcal{I} satisfying Σ

is the set $\text{ANS}(Q, \mathcal{I})$ composed by n -tuples \vec{o} such that

$$\text{ANS}(Q, \mathcal{I}) = \{ \vec{o} \mid \mathcal{I} \models \bigvee_j \exists \vec{y}_j. Q_j(\vec{o}, \vec{y}_j, \vec{c}_j) \}$$

Q_1 *is contained* in Q_2 with respect to a \mathcal{DLR} theory Σ

if and only if for every model \mathcal{I} satisfying Σ

$$\text{ANS}(Q_1, \mathcal{I}) \subseteq \text{ANS}(Q_2, \mathcal{I})$$

Semantics of Evaluation and Containment

The *evaluation* of a query Q of arity n given a \mathcal{DLR} theory Σ over a model \mathcal{I} satisfying Σ

is the set $\text{ANS}(Q, \mathcal{I})$ composed by n -tuples \vec{o} such that

$$\text{ANS}(Q, \mathcal{I}) = \{ \vec{o} \mid \mathcal{I} \models \bigvee_j \exists \vec{y}_j. Q_j(\vec{o}, \vec{y}_j, \vec{c}_j) \}$$

Q_1 *is contained* in Q_2 with respect to a \mathcal{DLR} theory Σ

if and only if for every model \mathcal{I} satisfying Σ

$$\text{ANS}(Q_1, \mathcal{I}) \subseteq \text{ANS}(Q_2, \mathcal{I})$$

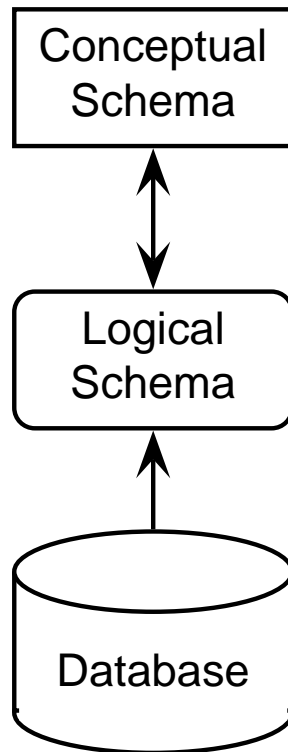
Containment of disjunctions of conjunctive queries under \mathcal{DLR} (\mathcal{DLR}_{US}) constraints is decidable in 2EXPTIME

The i●com tool for Intelligent Conceptual Modelling

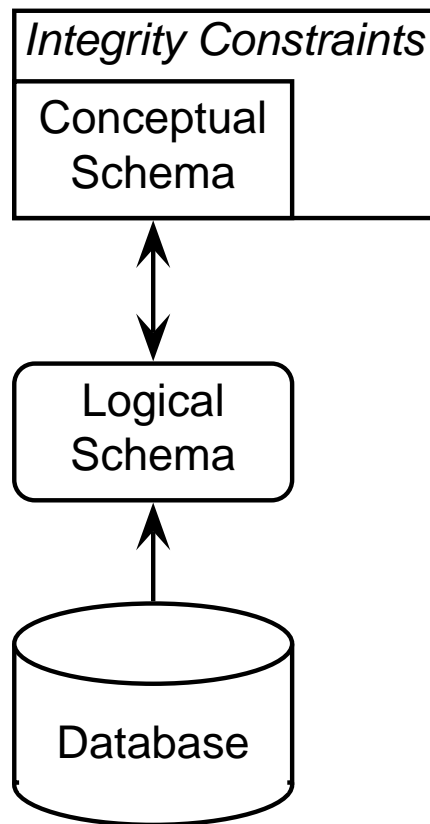
- i●com is an advanced CASE tool which allows the user to design *multiple* extended Entity-Relationship schemas or UML class diagrams with inter- and intra-schema *constraints*.
- Complete logical reasoning is employed by the tool to:
 - verify the specification,
 - infer implicit facts,
 - devise stricter constraints,
 - and manifest any local inconsistency.

<http://www.cs.man.ac.uk/~franconi/icom/>

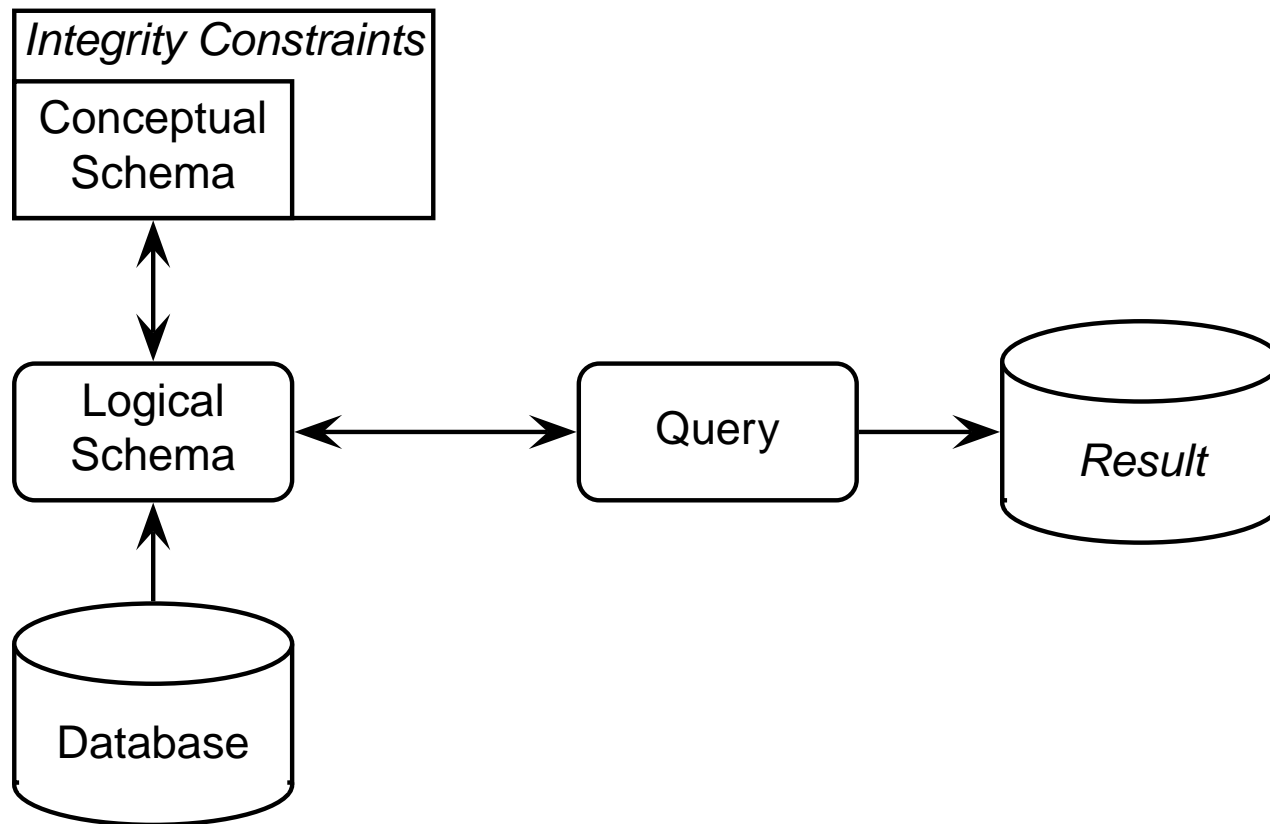
The role of logics in Information Systems



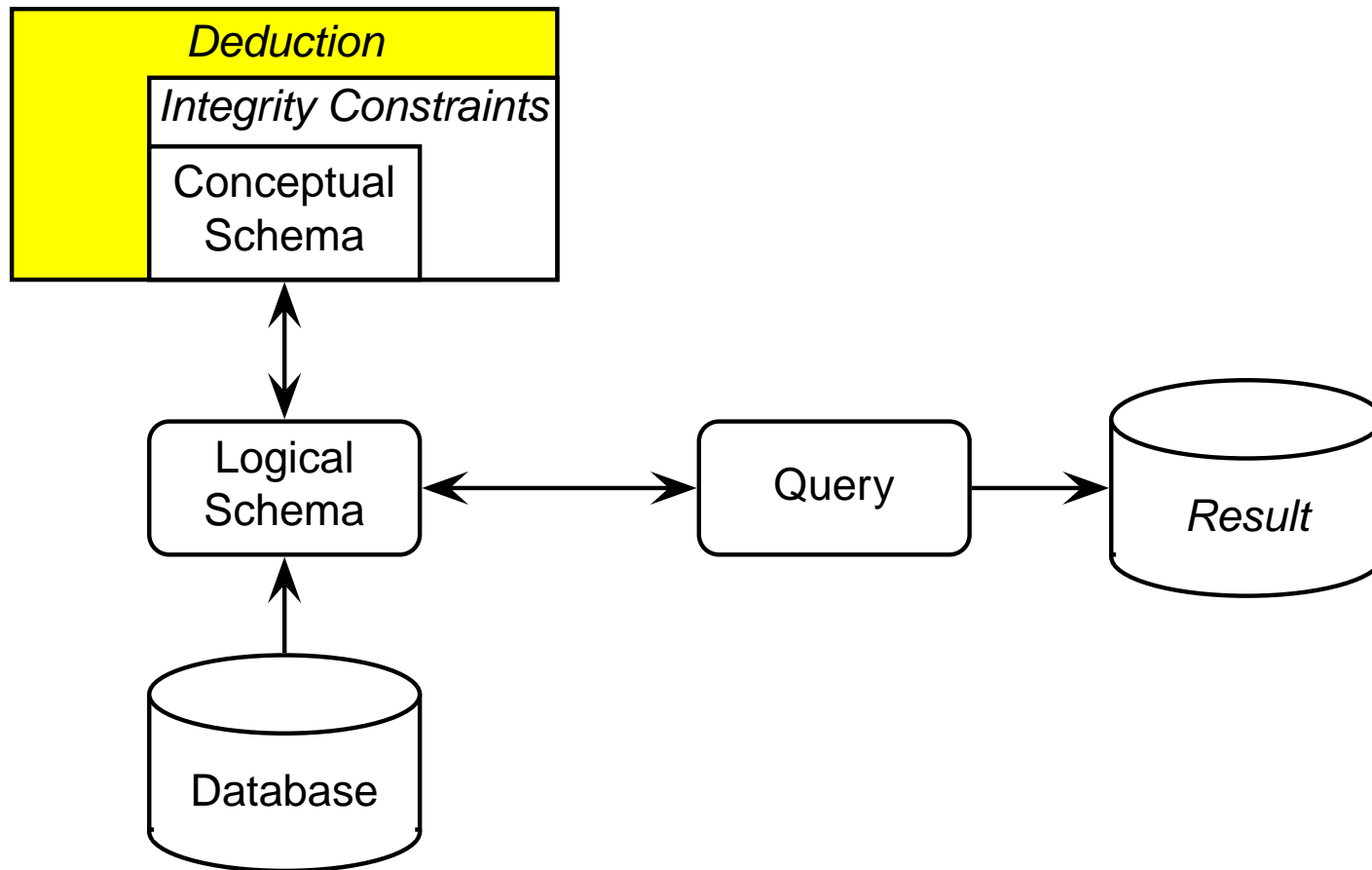
The role of logics in Information Systems



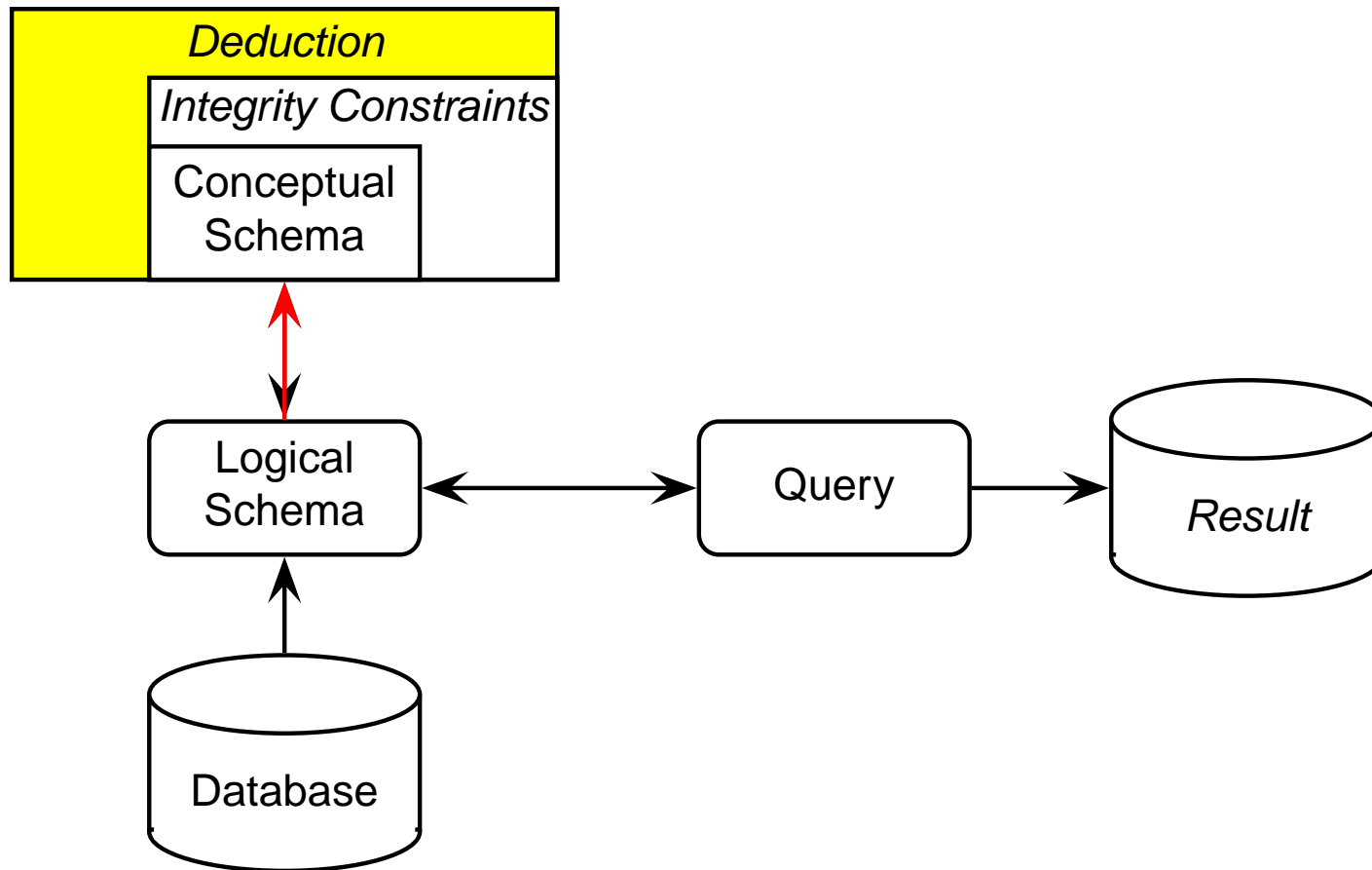
The role of logics in Information Systems



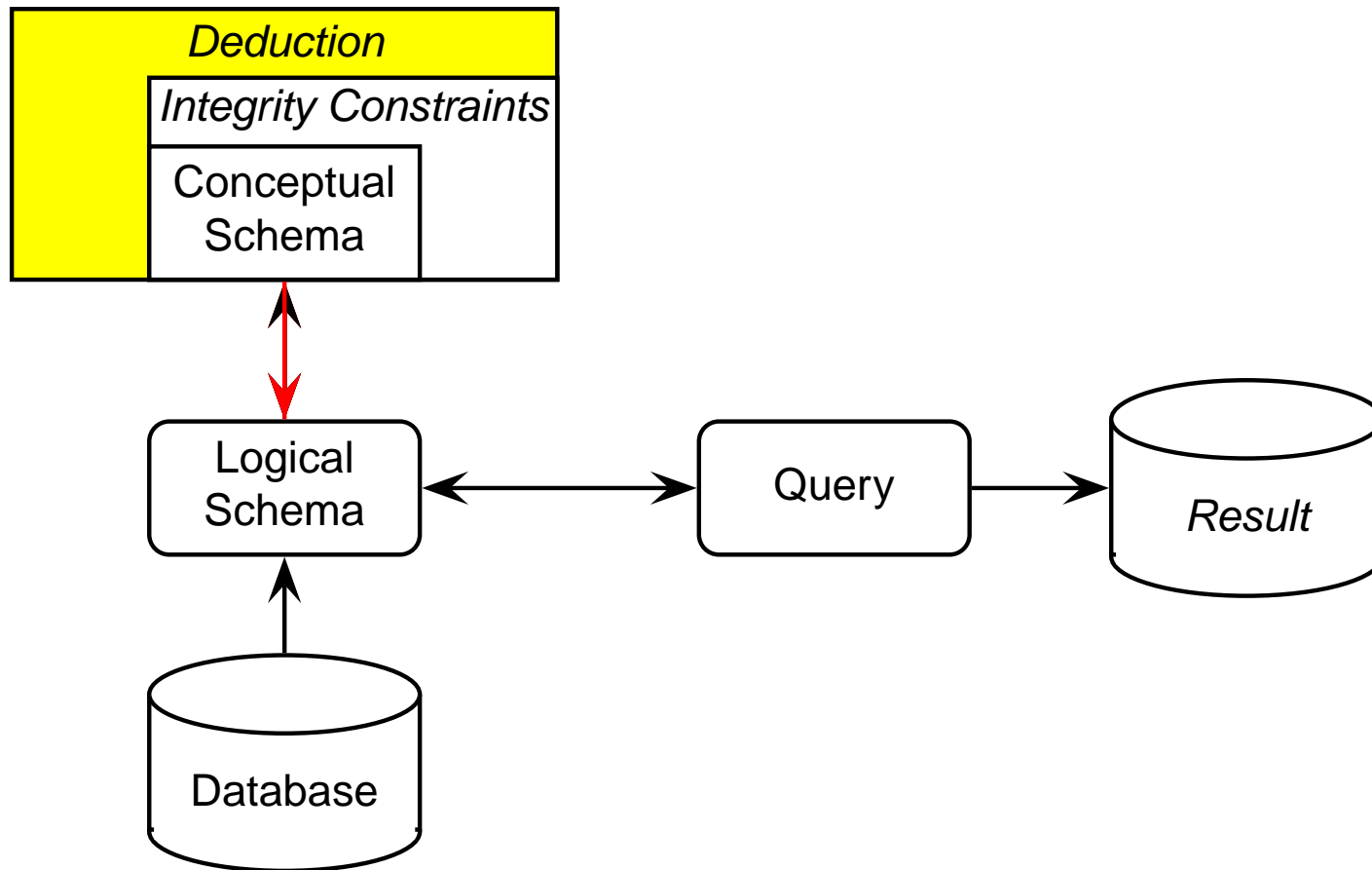
The role of logics in Information Systems



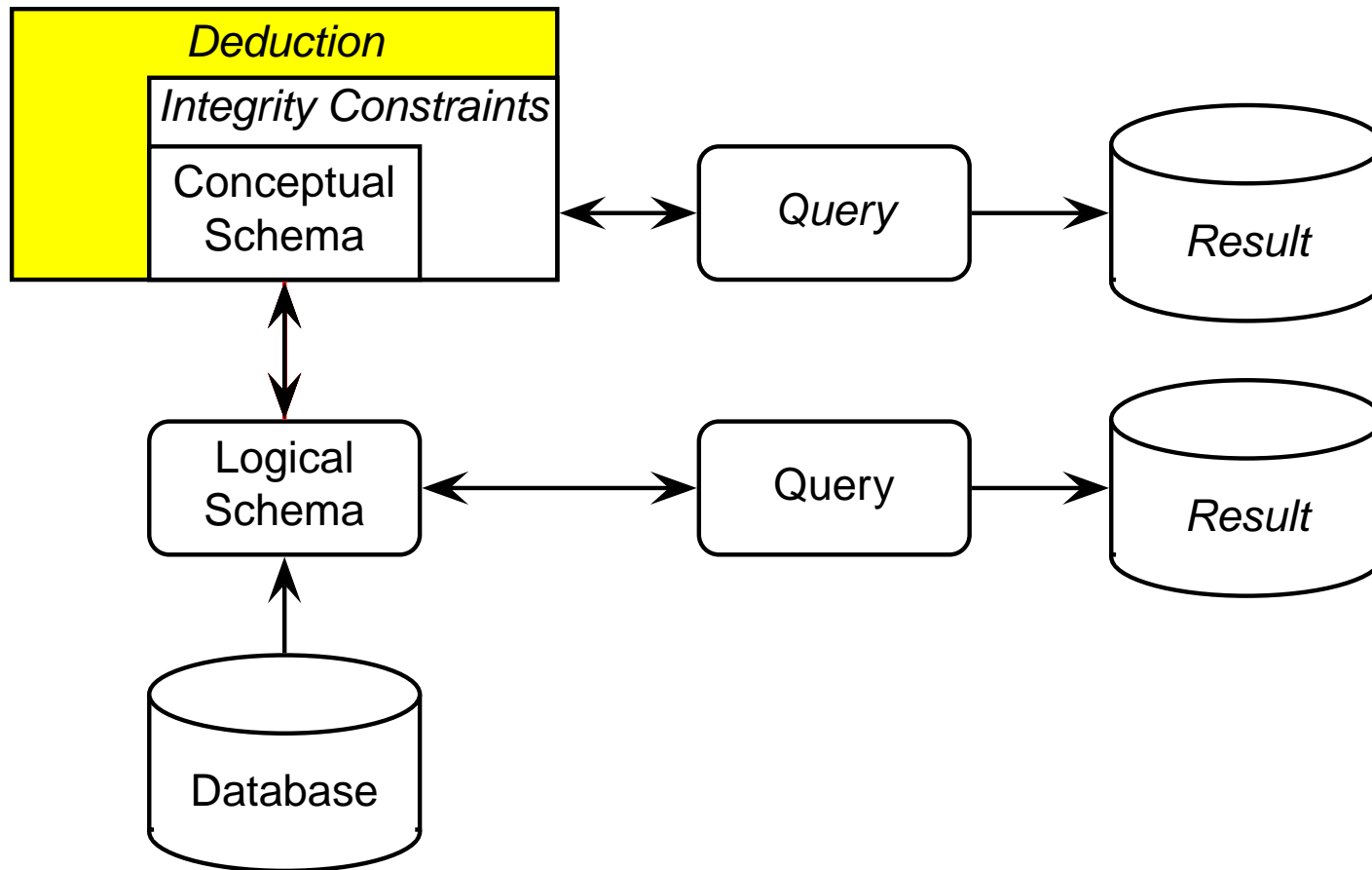
The role of logics in Information Systems



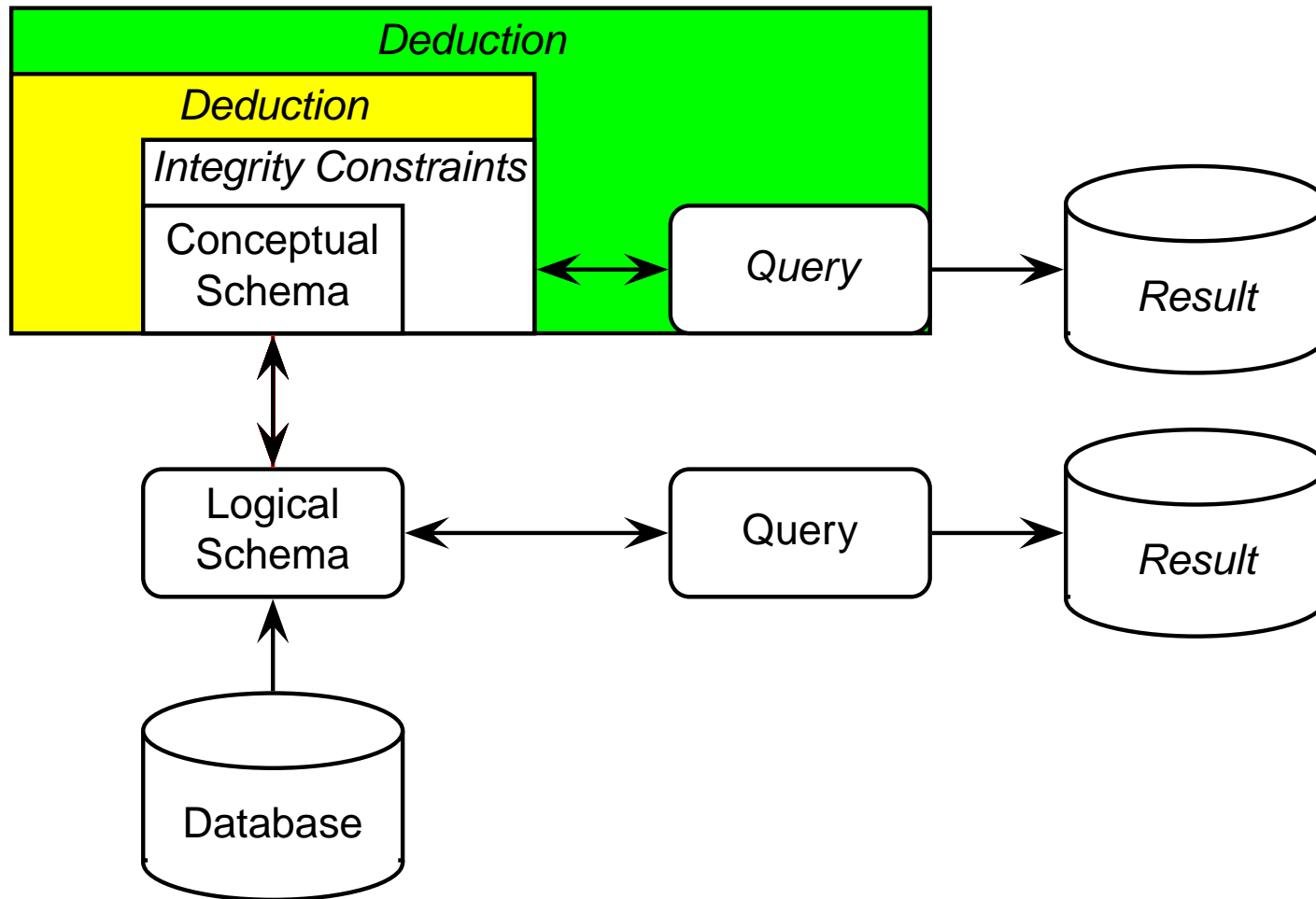
The role of logics in Information Systems



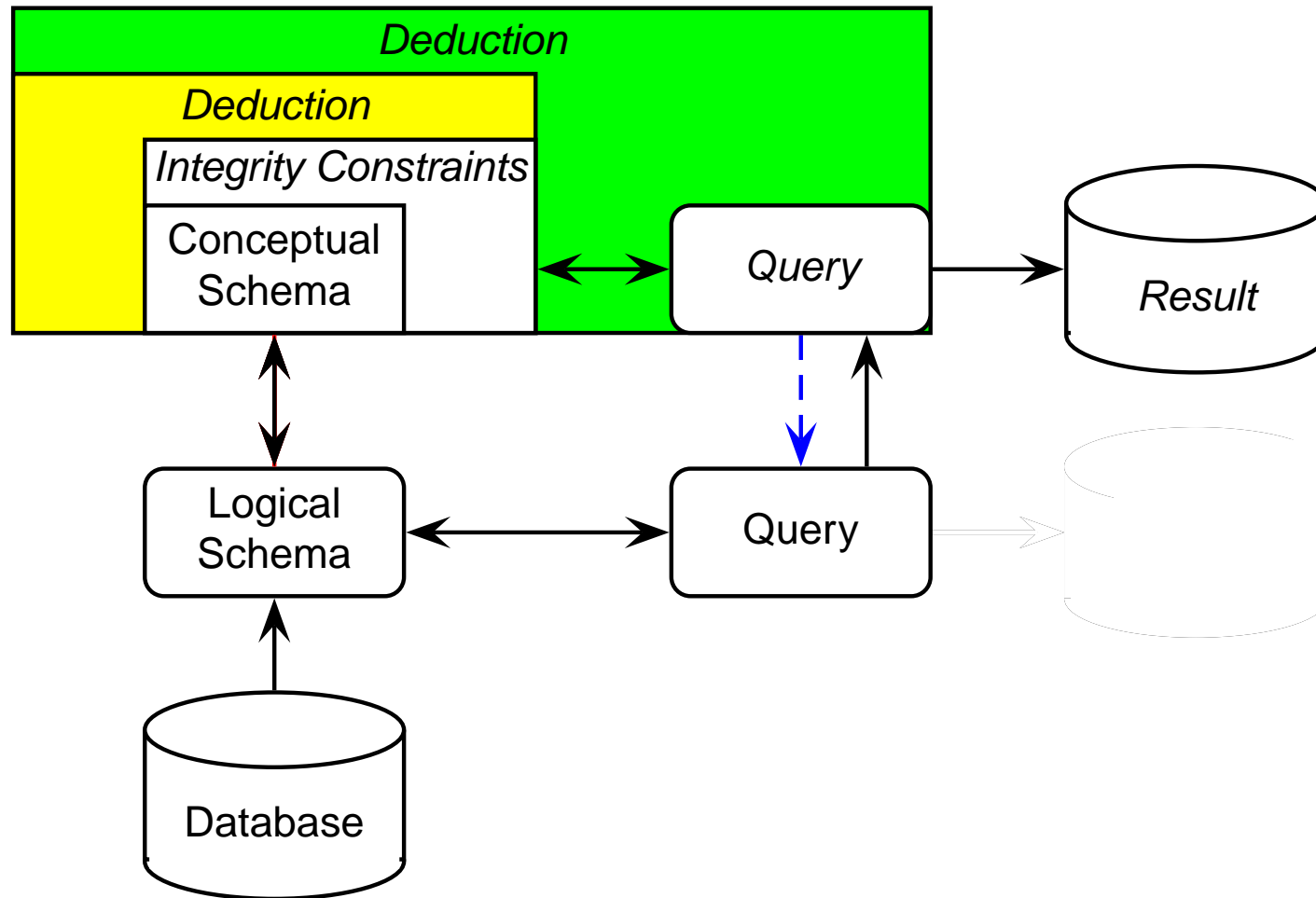
The role of logics in Information Systems



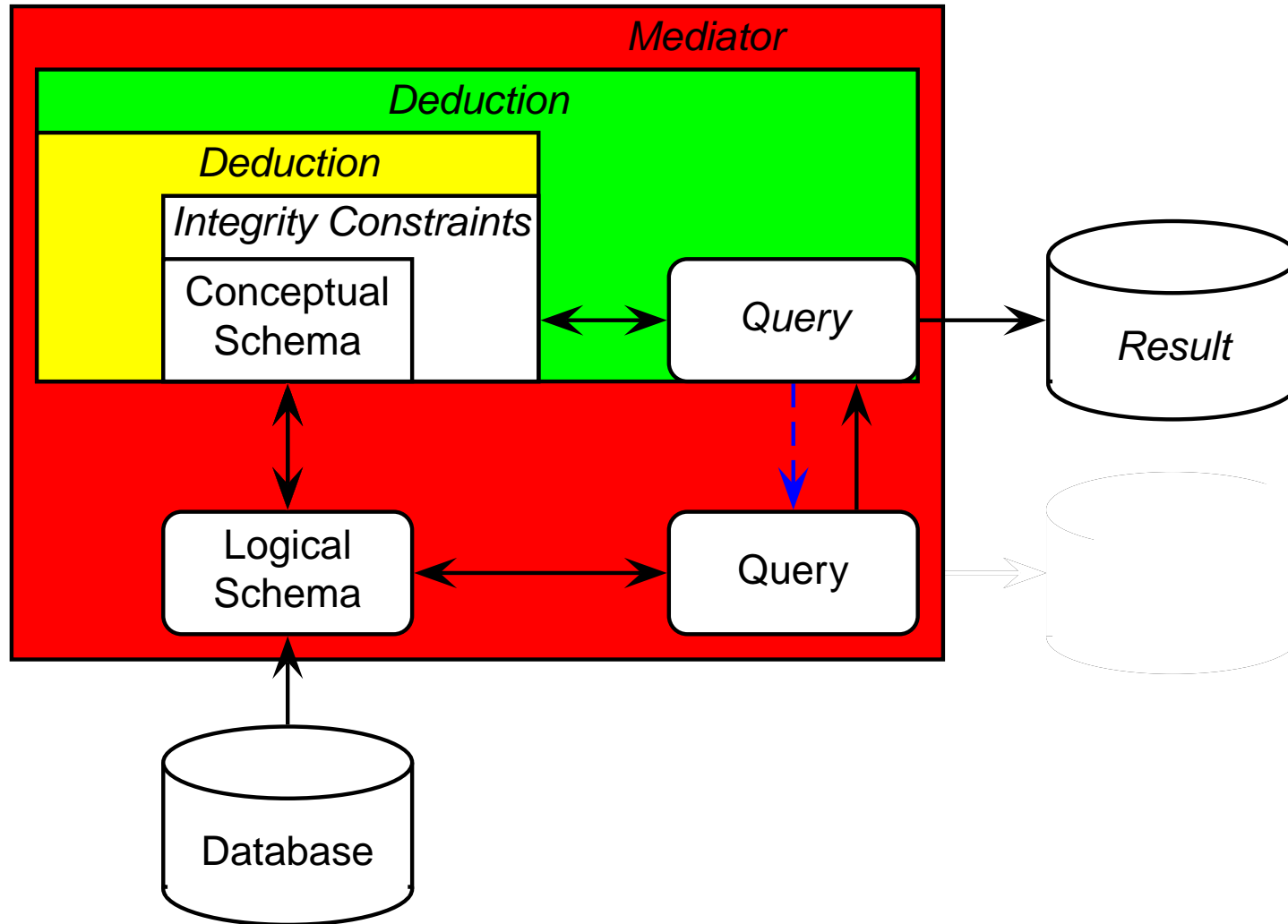
The role of logics in Information Systems



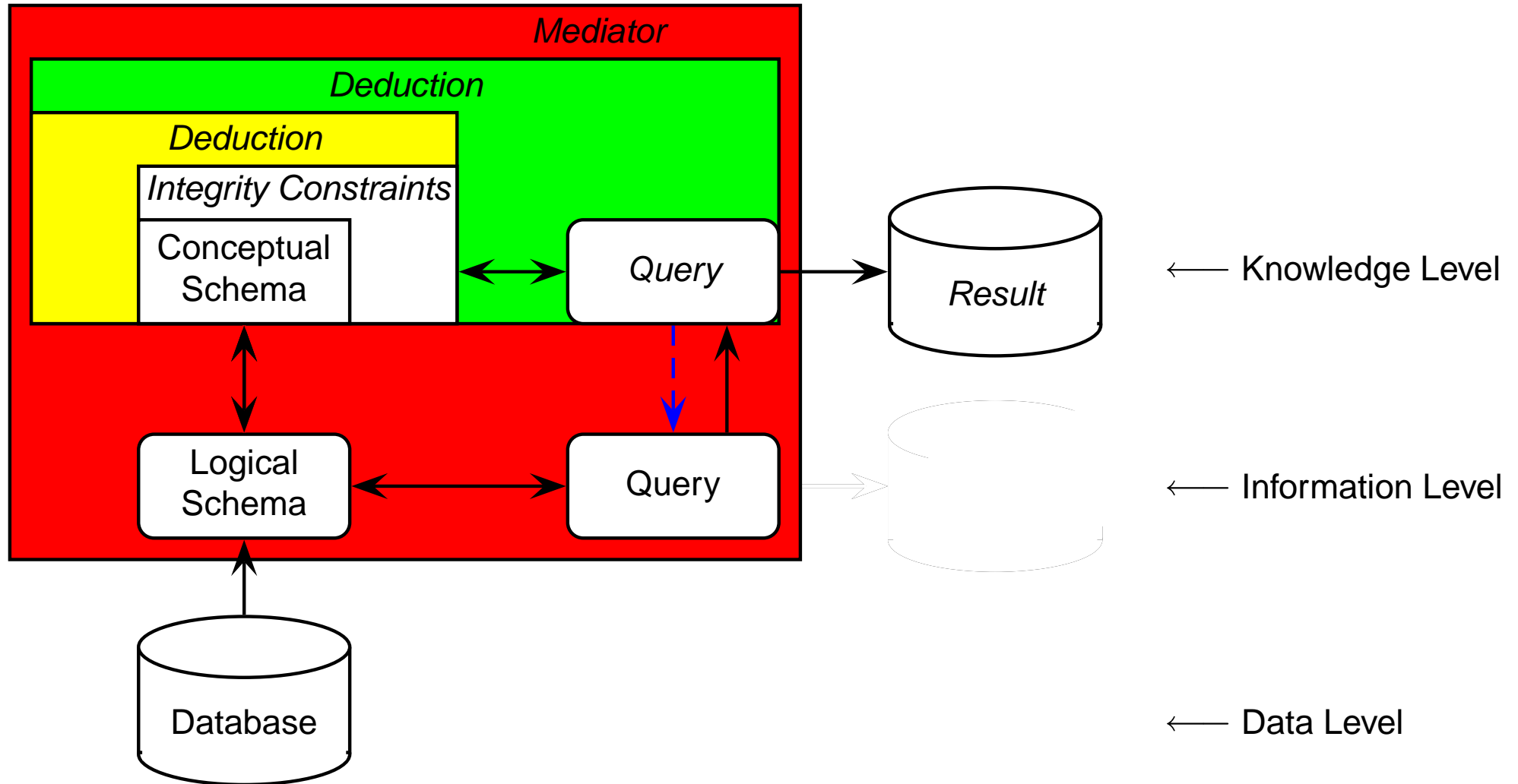
The role of logics in Information Systems



The role of logics in ~~Information Systems~~ a Mediator



The role of logics in ~~Information Systems~~ a Mediator



A Relational Database

CompanyEmployee/2; CompanyProject/3

CompanyEmployee	
name	project
john	esprit-dwq
...	...

CompanyProject		
project	manager	department
esprit-dwq	enrico	cs-uman
...

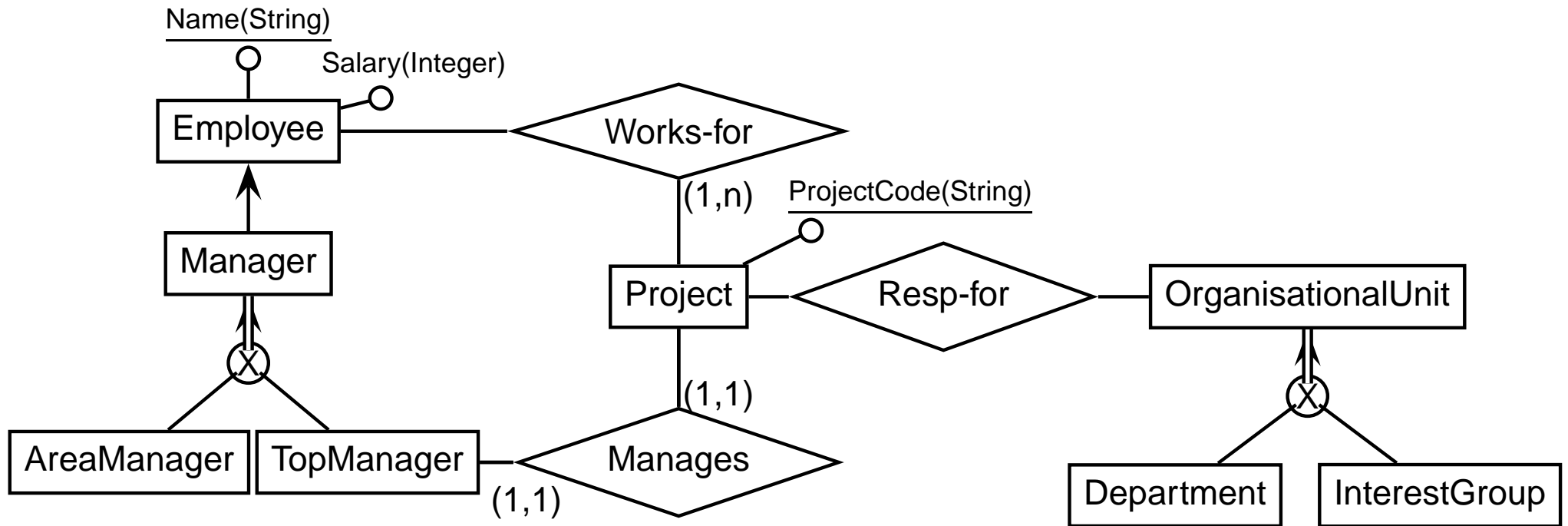
Query = “Tell me the projects in which John works, and their managers and departments.”

Query \equiv

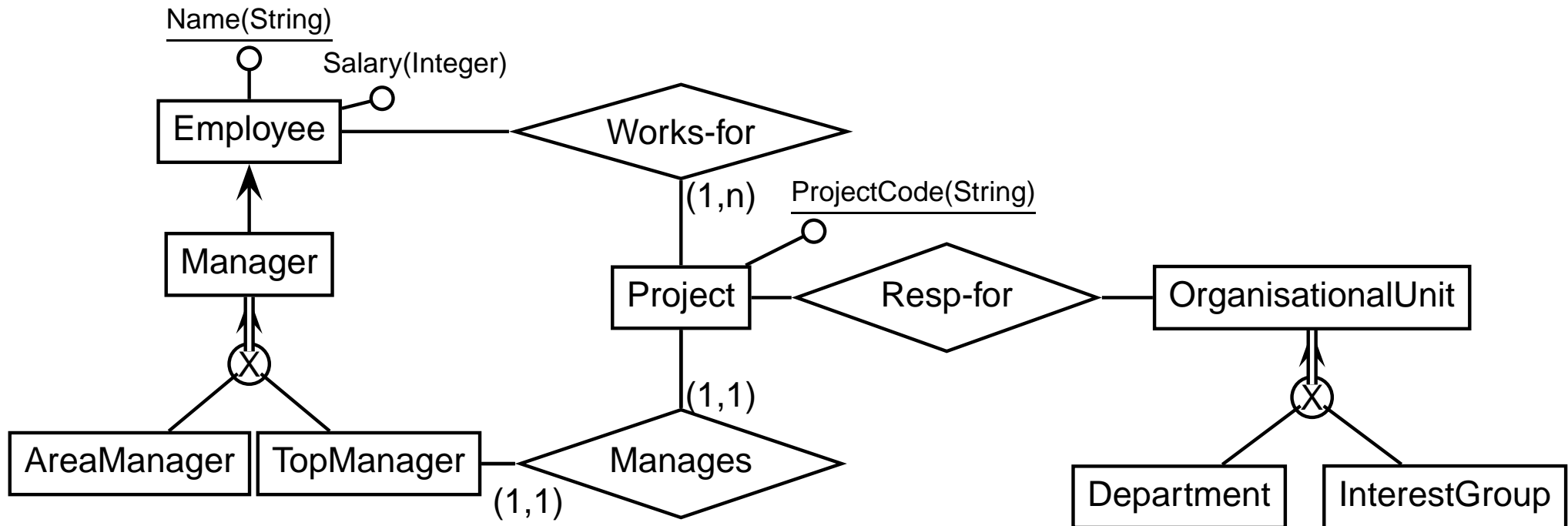
$\pi_{\text{project, manager, dept.}} \sigma_{\text{name=john}} (\text{CompanyEmployee} \bowtie_{\text{project}} \text{CompanyProject})$

$\text{Query}(x, y, z) \Leftrightarrow \text{CompanyEmployee}(\text{john}, x) \wedge \text{CompanyProject}(x, y, z)$

Constraints from the Conceptual Schema



Constraints from the Conceptual Schema



$\text{Works-for} \sqsubseteq \text{emp}/2 : \text{Employee} \sqcap \text{act}/2 : \text{Project}$

$\text{Manages} \sqsubseteq \text{man}/2 : \text{TopManager} \sqcap \text{prj}/2 : \text{Project}$

$\text{Employee} \sqsubseteq \exists^{=1}[\text{worker}](\text{Name} \sqcap \text{thename}/2 : \text{String}) \sqcap \exists^{=1}[\text{payee}](\text{Salary} \sqcap \text{amount}/2 : \text{Integer})$

$\top \sqsubseteq \exists^{\leq 1}[\text{thename}](\text{Name} \sqcap \text{worker}/2 : \text{Employee})$

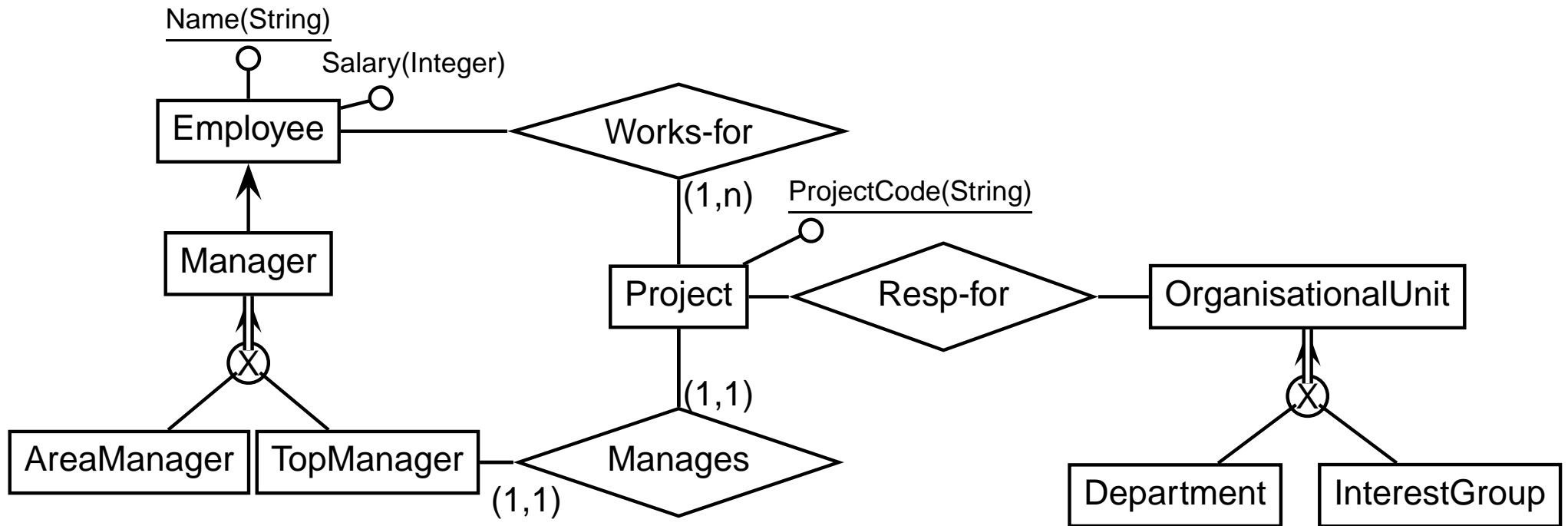
$\text{Manager} \sqsubseteq \text{Employee} \sqcap (\text{AreaManager} \sqcup \text{TopManager})$

$\text{AreaManager} \sqsubseteq \text{Manager} \sqcap \neg \text{TopManager}$

$\text{TopManager} \sqsubseteq \text{Manager} \sqcap \exists^{=1}[\text{man}]\text{Manages}$

$\text{Project} \sqsubseteq \exists^{\geq 1}[\text{act}]\text{Works-for} \sqcap \exists^{=1}[\text{prj}]\text{Manages}$
 \dots

Integrity Constraints and Logical Implication

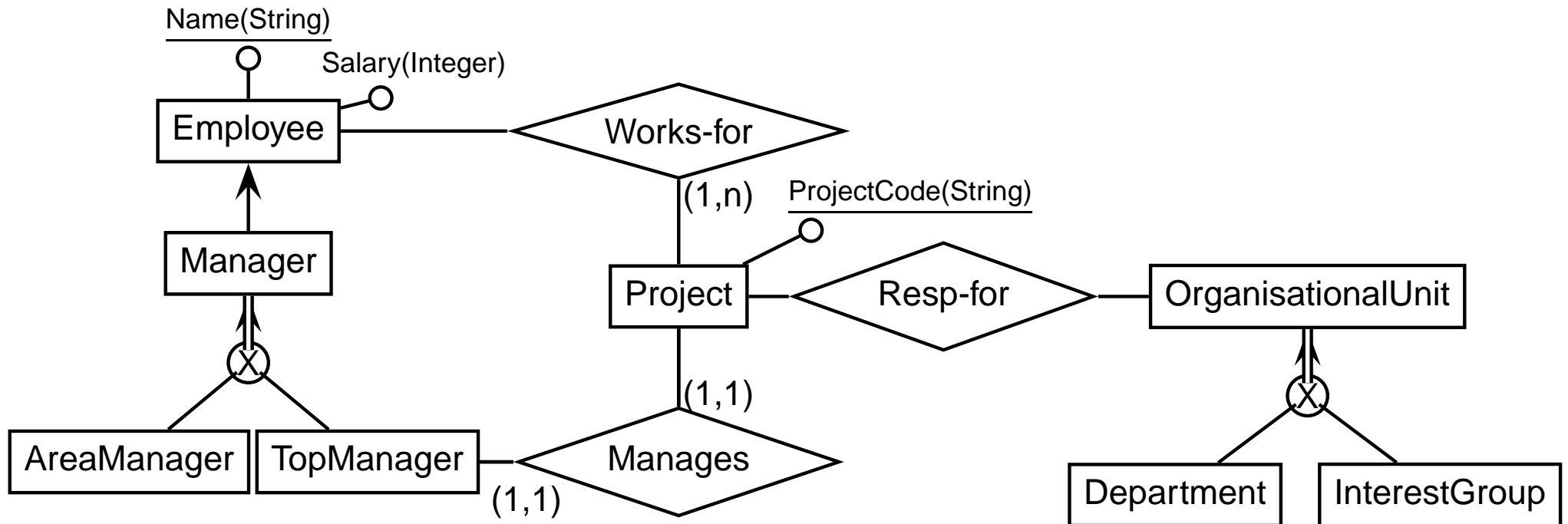


Managers are employees who do not work for a project (she/he just manages it):

$$\text{Employee} \sqcap \neg(\exists^{\geq 1}[\text{emp}]\text{Works-for}) \sqsubseteq \text{Manager}$$

$$\text{Manager} \sqsubseteq \neg(\exists^{\geq 1}[\text{emp}]\text{Works-for})$$

Integrity Constraints and Logical Implication



Managers are employees who do not work for a project (she/he just manages it):

$$\text{Employee} \sqcap \neg(\exists^{\geq 1}[\text{emp}]\text{Works-for}) \sqsubseteq \text{Manager}$$

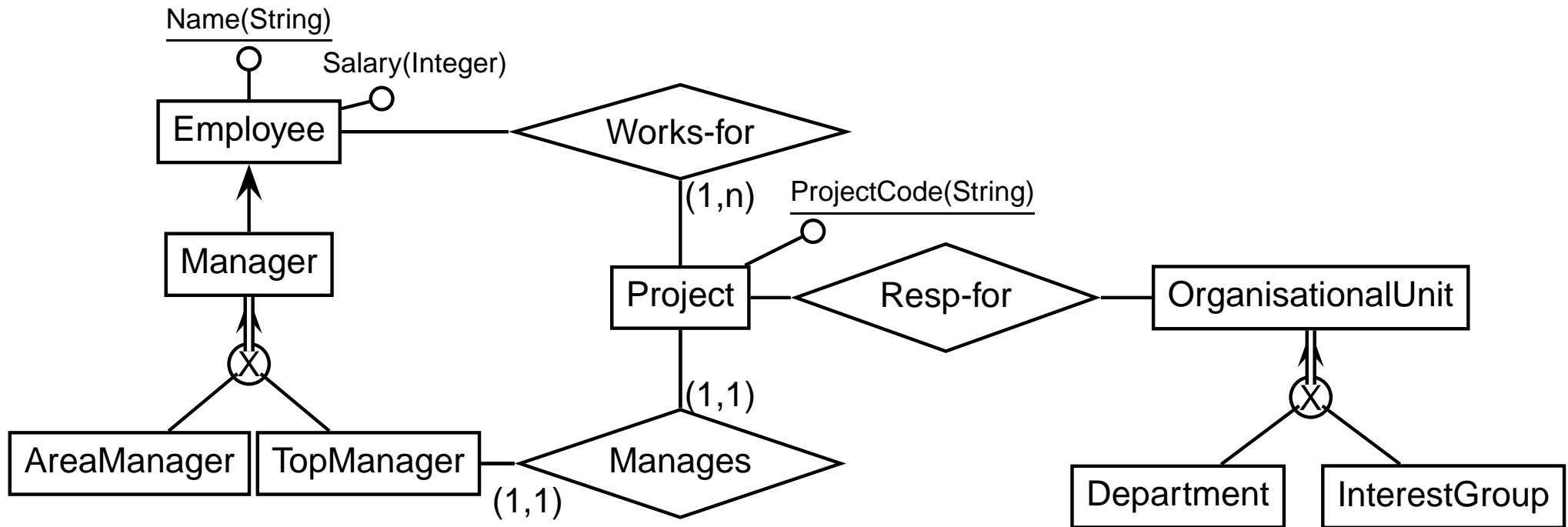
$$\text{Manager} \sqsubseteq \neg(\exists^{\geq 1}[\text{emp}]\text{Works-for})$$

\leadsto For every project, there is at least one employee who is not a manager:

$$\Sigma \models \text{Project} \sqsubseteq \exists^{\geq 1}[\text{act}](\text{Works-for} \sqcap \text{emp} : \neg \text{Manager})$$

Querying the Virtual Database (local-as-view)

$$Q(x, y, z) \Leftarrow \text{Project}(x) \wedge \text{Works-for}(\text{john}, x) \wedge \text{TopManager}(y) \wedge \text{Manages}(y, x) \wedge \neg \text{InterestGroup}(z) \wedge \text{Resp-for}(z, x).$$

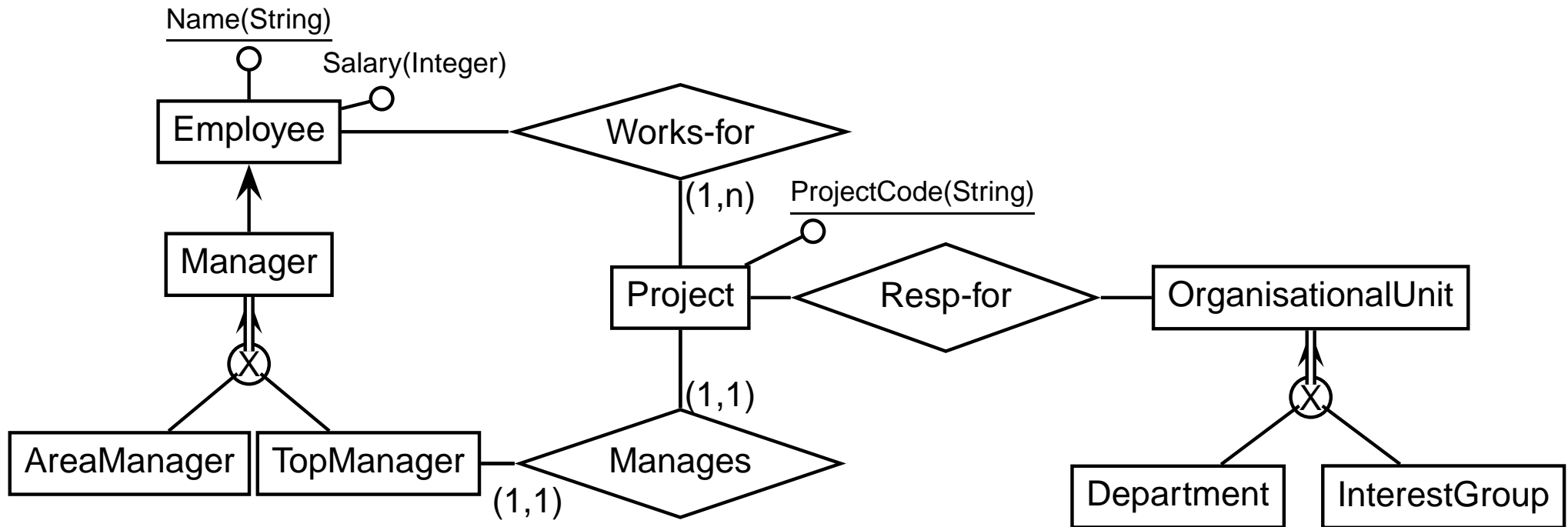


$$\text{CompanyEmployee}(x, y) \Leftarrow \text{Employee}(x) \wedge \text{Project}(y) \wedge \text{Works-for}(x, y).$$

$$\text{CompanyProject}(x, y, z) \Leftarrow \text{Project}(x) \wedge \text{Manager}(y) \wedge \text{Department}(z) \wedge \text{Manages}(y, x) \wedge \text{Resp-for}(z, x).$$

Querying the Virtual Database (local-as-view)

$$Q(x, y, z) \Leftarrow \text{Project}(x) \wedge \text{Works-for}(\text{john}, x) \wedge \text{TopManager}(y) \wedge \text{Manages}(y, x) \wedge \neg \text{InterestGroup}(z) \wedge \text{Resp-for}(z, x).$$

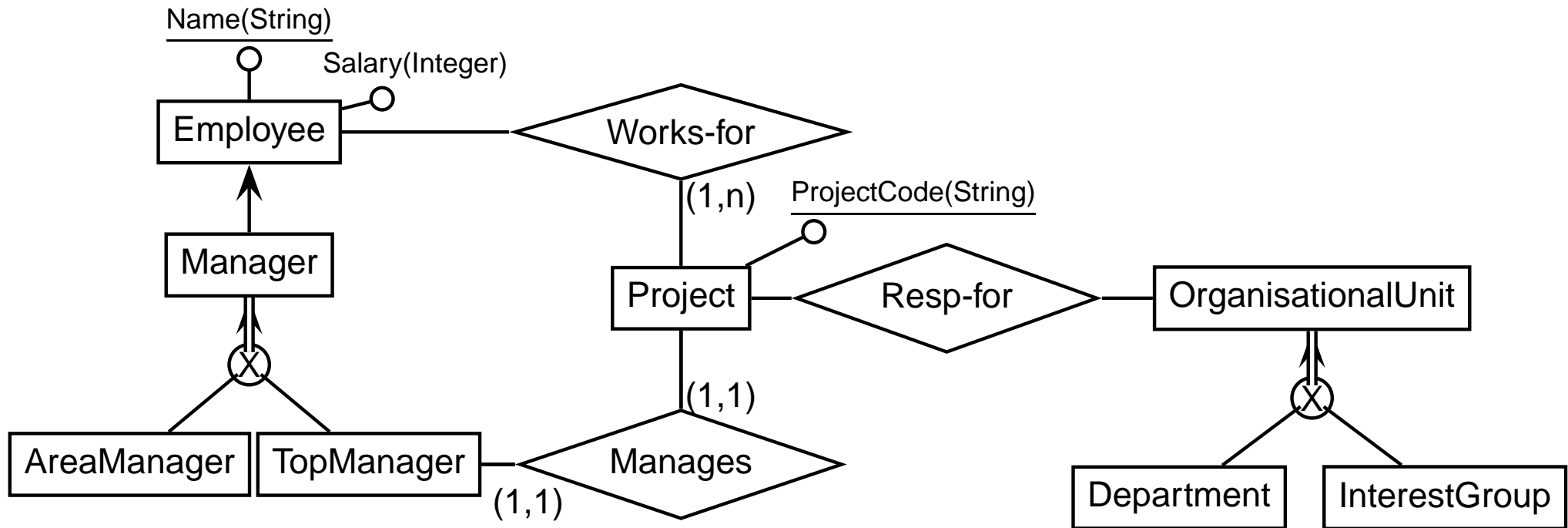


$$\text{CompanyEmployee}(x, y) \Leftarrow \text{Employee}(x) \wedge \text{Project}(y) \wedge \text{Works-for}(x, y).$$

$$\text{CompanyProject}(x, y, z) \Leftarrow \text{Project}(x) \wedge \text{Manager}(y) \wedge \text{Department}(z) \wedge \text{Manages}(y, x) \wedge \text{Resp-for}(z, x).$$

$$\leadsto Q(x, y, z) \Leftarrow \text{CompanyEmployee}(\text{john}, x) \wedge \text{CompanyProject}(x, y, z)$$

Querying the Virtual Database (global-as-view)

$$Q(x, y, z) \Leftarrow \text{Project}(x) \wedge \text{Works-for}(\text{john}, x) \wedge \text{TopManager}(y) \wedge \text{Manages}(y, x) \wedge \neg \text{InterestGroup}(z) \wedge \text{Resp-for}(z, x).$$


$$\text{Project}(x) \Leftarrow \text{CompanyEmployee}(y, x) \cup \text{CompanyProject}(x, y, z)$$

$$\text{Works-for}(x, y) \Leftarrow \text{CompanyEmployee}(x, y)$$

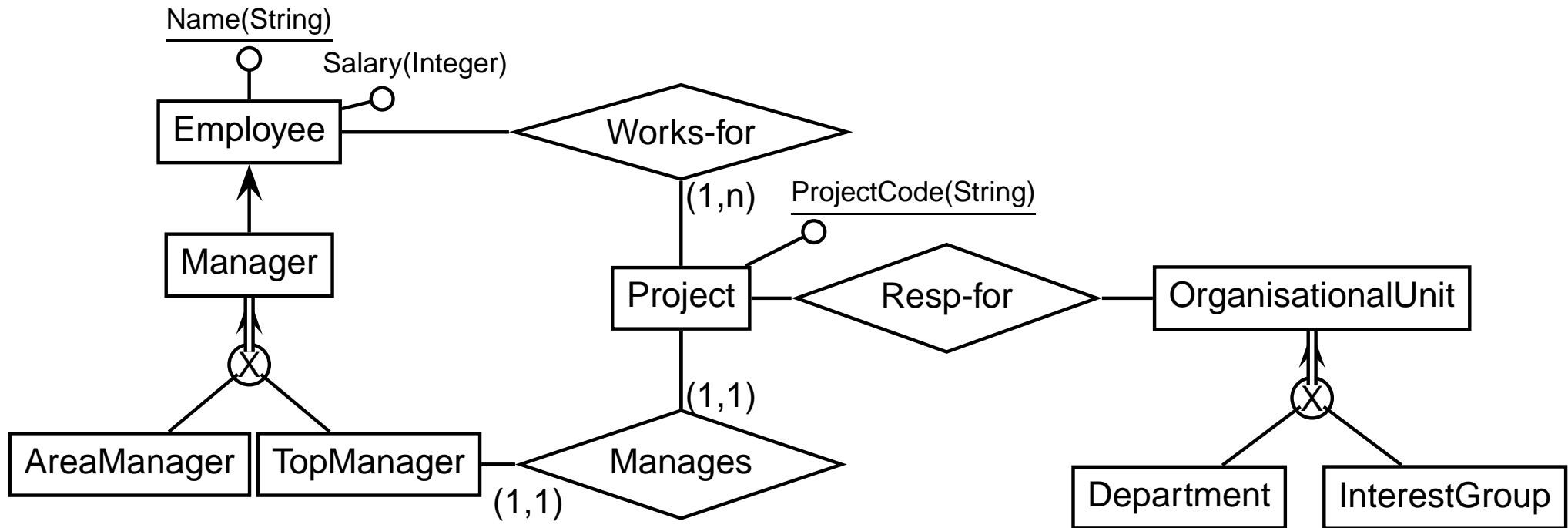
$$\text{TopManager}(x) \Leftarrow \text{CompanyProject}(y, x, z)$$

$$\text{Manages}(x, y) \Leftarrow \text{CompanyProject}(y, x, z)$$

...

Querying the Virtual Database (global-as-view)

$$Q(x, y, z) \Leftarrow \text{Project}(x) \wedge \text{Works-for}(\text{john}, x) \wedge \text{TopManager}(y) \wedge \text{Manages}(y, x) \wedge \neg \text{InterestGroup}(z) \wedge \text{Resp-for}(z, x).$$



$$\text{Project}(x) \Leftarrow \text{CompanyEmployee}(y, x) \cup \text{CompanyProject}(x, y, z)$$

$$\text{Works-for}(x, y) \Leftarrow \text{CompanyEmployee}(x, y)$$

$$\text{TopManager}(x) \Leftarrow \text{CompanyProject}(y, x, z)$$

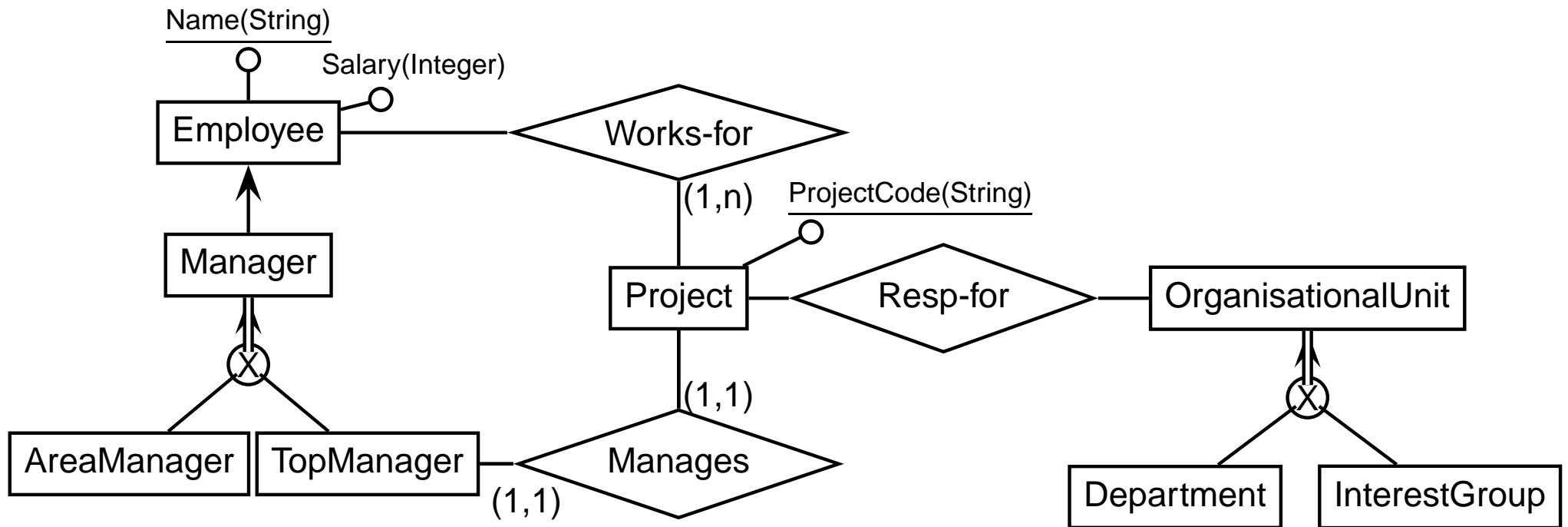
$$\text{Manages}(x, y) \Leftarrow \text{CompanyProject}(y, x, z)$$

...

$$\rightsquigarrow Q(x, y, z) \Leftarrow \text{CompanyEmployee}(\text{john}, x) \wedge \text{CompanyProject}(x, y, z)$$

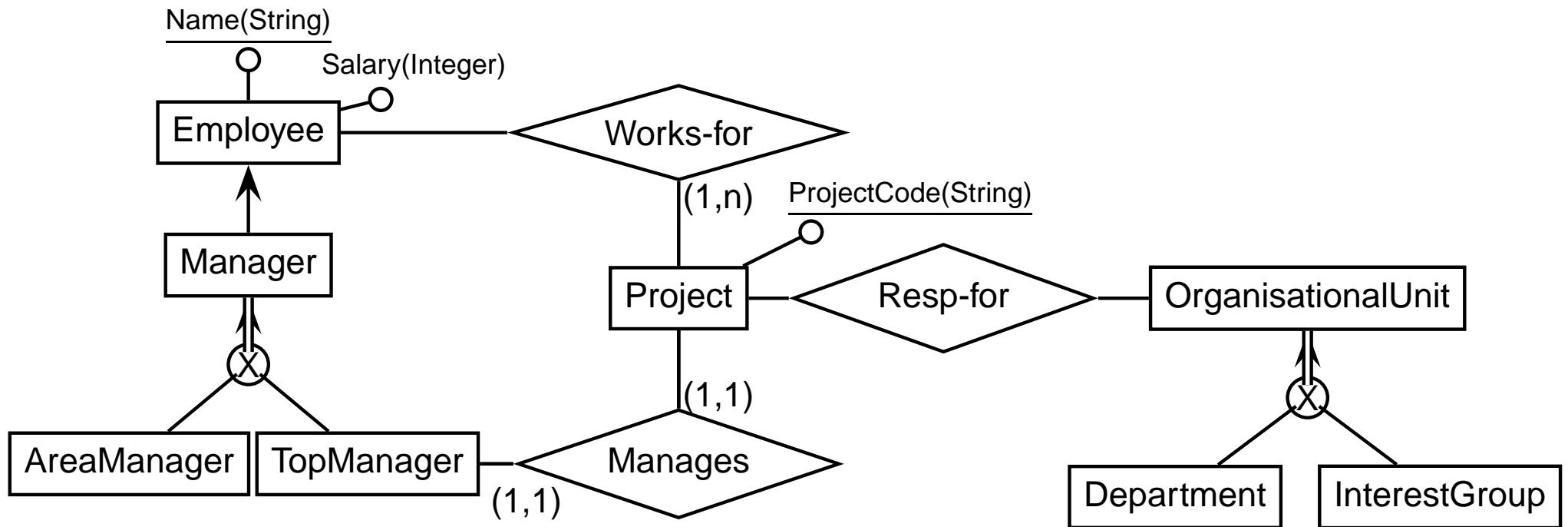
Reasoning on Queries

$Q(x, y) \Leftarrow \text{Employee}(x) \wedge \text{Works-for}(x, y) \wedge \text{Manages}(x, y).$



Reasoning on Queries

$$Q(x, y) \Leftarrow \text{Employee}(x) \wedge \text{Works-for}(x, y) \wedge \text{Manages}(x, y).$$



~> INCONSISTENT QUERY!