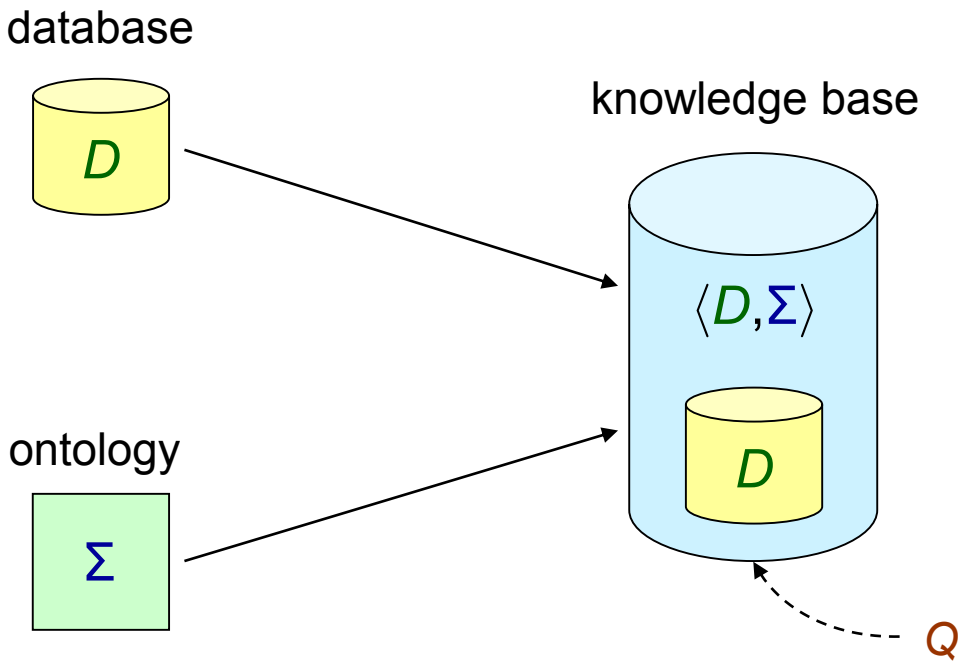
Consistent Query Answering in OBDA

Ontology-Based Query Answering (OBQA)



existential rules

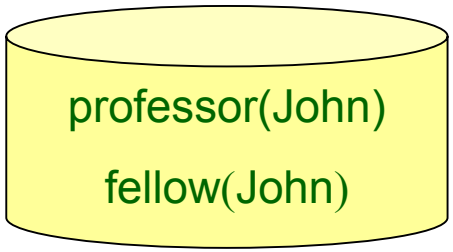
$\forall \mathbf{x} \forall \mathbf{y}\ (\varphi(\mathbf{x},\mathbf{y}) \rightarrow \exists \mathbf{z}\ \psi(\mathbf{x},\mathbf{z}))$

conjunctive queries

$Q(\mathbf{x})\ \text{:-}\ R_1(\mathbf{v_1}),\ldots,R_m(\mathbf{v_m})$

A Simple Example

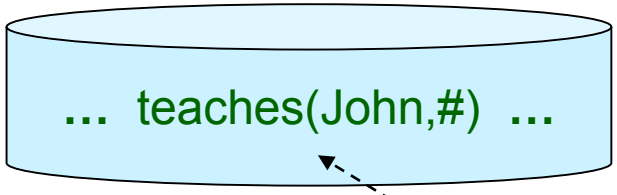
D =



Σ =

$\forall x\ (professor(x) \rightarrow \exists y\ (faculty(x) \wedge teaches(x,y)))$

$\forall x\ (fellow(x) \rightarrow faculty(x))$

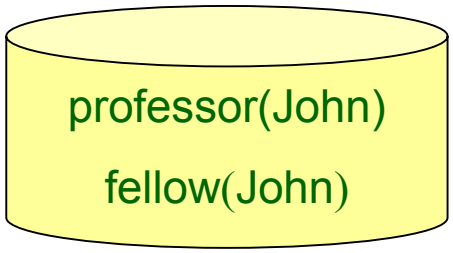


{John →John, x →#}

Q :- (teaches(John,x)) ✓

A Simple Example

D =



Σ =

$\forall x\ (professor(x) \rightarrow \exists y\ (faculty(x) \wedge teaches(x,y)))$

$\forall x\ (fellow(x) \rightarrow faculty(x))$

$\forall x\ (professor(x) \wedge fellow(x) \rightarrow \bot)$

no model $\Rightarrow$ every query is entailed

Handling Data Inconsistencies

- The data are likely to be inconsistent with the ontology
- Standard semantics fails: everything is inferred - not meaningful answers
- Two approaches to inconsistency-handling:
  - Resolve the inconsistencies - ideal, but not always possible
  - Live with the inconsistencies - inconsistency-tolerant semantics

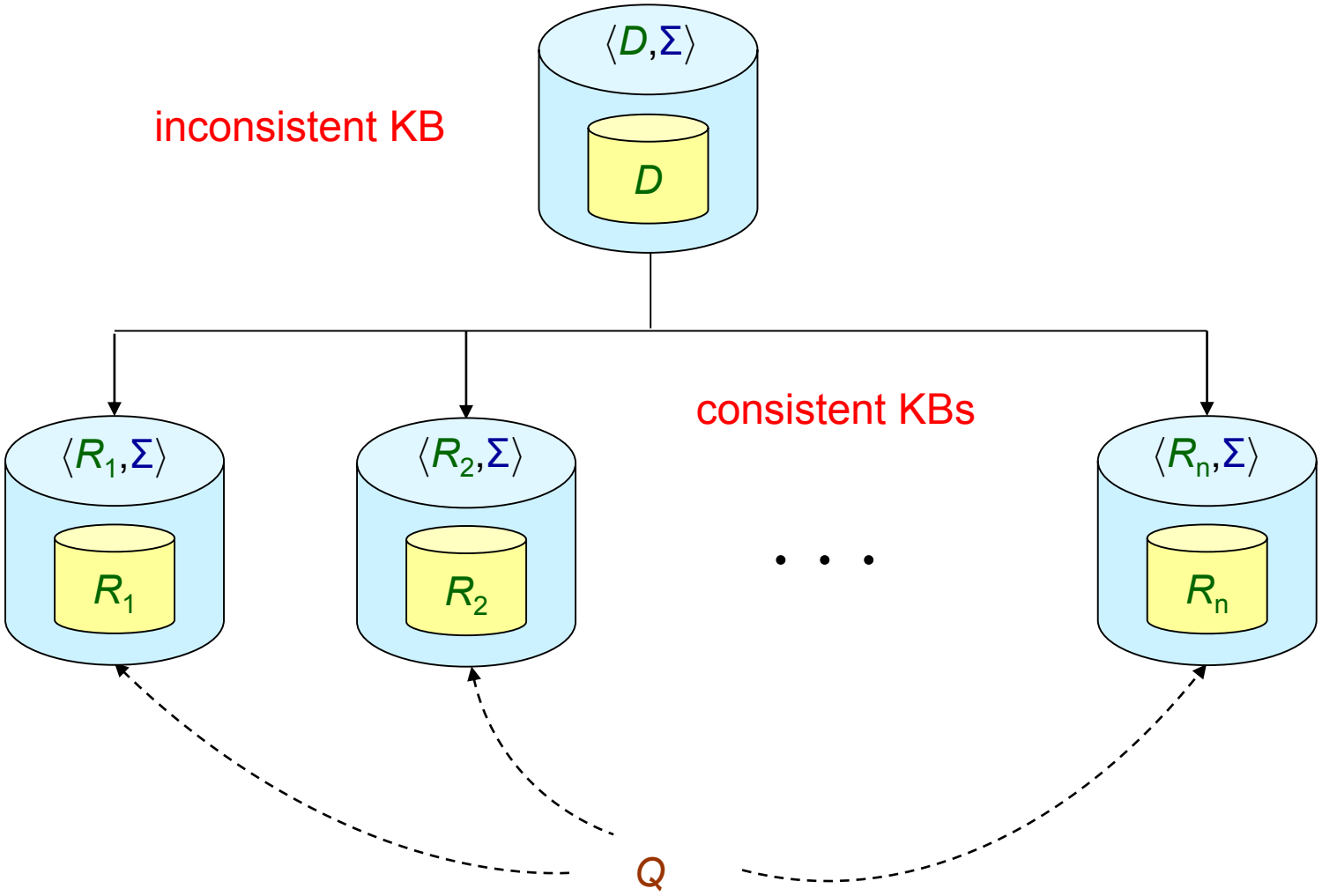
ABox Repair (AR) Semantics

- Standard inconsistency-tolerant semantics

- **IDEA:** The query must be entailed by every database repair

$\subseteq$-maximal consistent subsets of the database

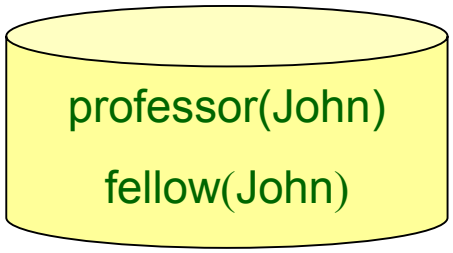
ABox Repair (AR) Semantics



AR-answers(Q, $\langle D,\Sigma \rangle$) = $\bigcap_{R \in \{R_1,\ldots,R_n\}}$ certain-answers(Q, $\langle R,\Sigma \rangle$)

ABox Repair (AR) Semantics: Example

D =



Σ =

$\forall x\ (professor(x) \rightarrow \exists y\ (faculty(x) \wedge teaches(x,y)))$

$\forall x\ (fellow(x) \rightarrow faculty(x))$

$\forall x\ (professor(x) \wedge fellow(x) \rightarrow \bot)$

Q :- faculty(John) ✓

R_1 = professor(John)

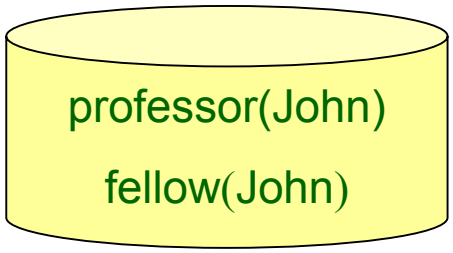
$() \in$ AR-answers(Q, $\langle R_1, \Sigma \rangle$)

R_2 = fellow(John)

$() \in$ AR-answers(Q, $\langle R_2, \Sigma \rangle$)

ABox Repair (AR) Semantics: Example

D =



Σ =

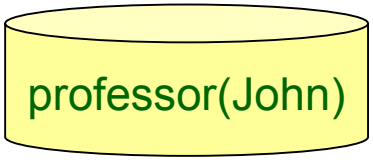
$\forall x\ (professor(x) \rightarrow \exists y\ (faculty(x) \wedge teaches(x,y)))$

$\forall x\ (fellow(x) \rightarrow faculty(x))$

$\forall x\ (professor(x) \wedge fellow(x) \rightarrow \bot)$

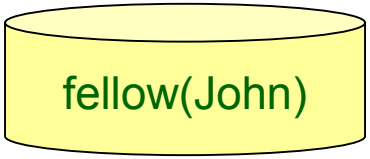
Q :- teaches(John,x) ✗

R_1 =



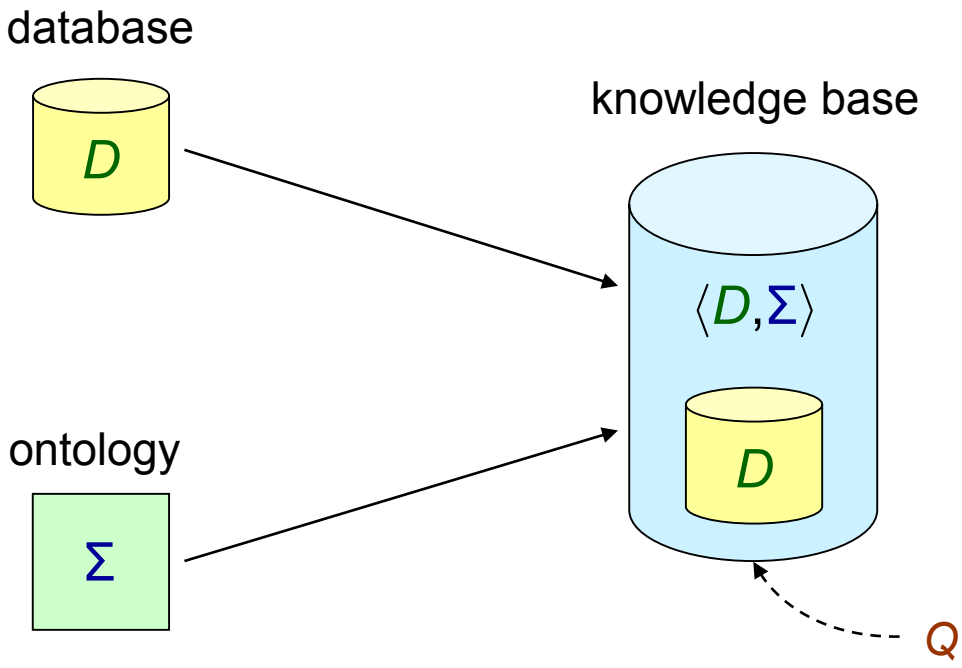
$() \in$ AR-answers(Q, $\langle R_1, \Sigma \rangle$)

R_2 =



$() \notin$ AR-answers(Q, $\langle R_2, \Sigma \rangle$)

Consistent Query Answering in OBQA



existential rules + negative constraints

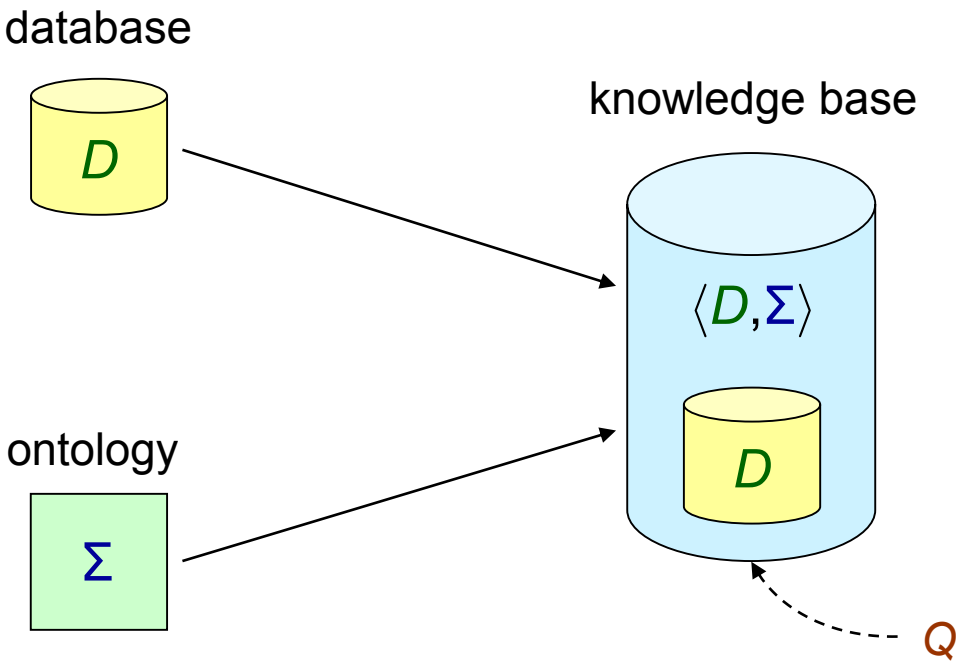
$\forall \mathbf{x} \forall \mathbf{y}\ (\varphi(\mathbf{x},\mathbf{y}) \rightarrow \exists \mathbf{z}\ \psi(\mathbf{x},\mathbf{z}))$

$\forall \mathbf{x}\ (\varphi(\mathbf{x}) \rightarrow \bot)$

conjunctive queries

$Q(\mathbf{x})\ \text{:-}\ R_1(\mathbf{v_1}),\ldots,R_m(\mathbf{v_m})$

Consistent Query Answering in OBQA



AR-answers(Q, $\langle D,\Sigma \rangle$) = $\bigcap_{R \in \text{drep}(D,\Sigma)}$ certain-answers(Q, $\langle R,\Sigma \rangle$)

$\{D' \mid D \supseteq D'$, models$(D' \wedge \Sigma) \neq \emptyset$, there is no $\alpha \in D$ such that models$(D' \cup \{\alpha\} \wedge \Sigma) \neq \emptyset\}$

Consistent Query Answering in OBQA

Guess and check algorithm (for the complement of the problem)

Input: D, Σ, $Q(x)$, tuple $\mathbf{t}$

1. Guess $R \subseteq D$ - a possible repair

2. Verify that R is a repair, i.e., $\langle R,\Sigma\rangle$ is consistent and R is $\subseteq$-maximal

3. Verify that $\langle R,\Sigma\rangle$ does not entail $Q(\mathbf{t})$

we exploit classical query answering

2.1. Check that for every $\forall \mathbf{x}\ (\varphi(\mathbf{x}) \rightarrow \bot) \in \Sigma$, $\langle R,\Sigma\rangle$ does not entail $Q :- \varphi(\mathbf{x})$

2.2. Check that for every $\alpha \in D \setminus R$, there exists $\forall \mathbf{x}\ (\varphi(\mathbf{x}) \rightarrow \bot) \in \Sigma$, such that $\langle R \cup \{\alpha\},\Sigma\rangle$ entails $Q :- \varphi(\mathbf{x})$

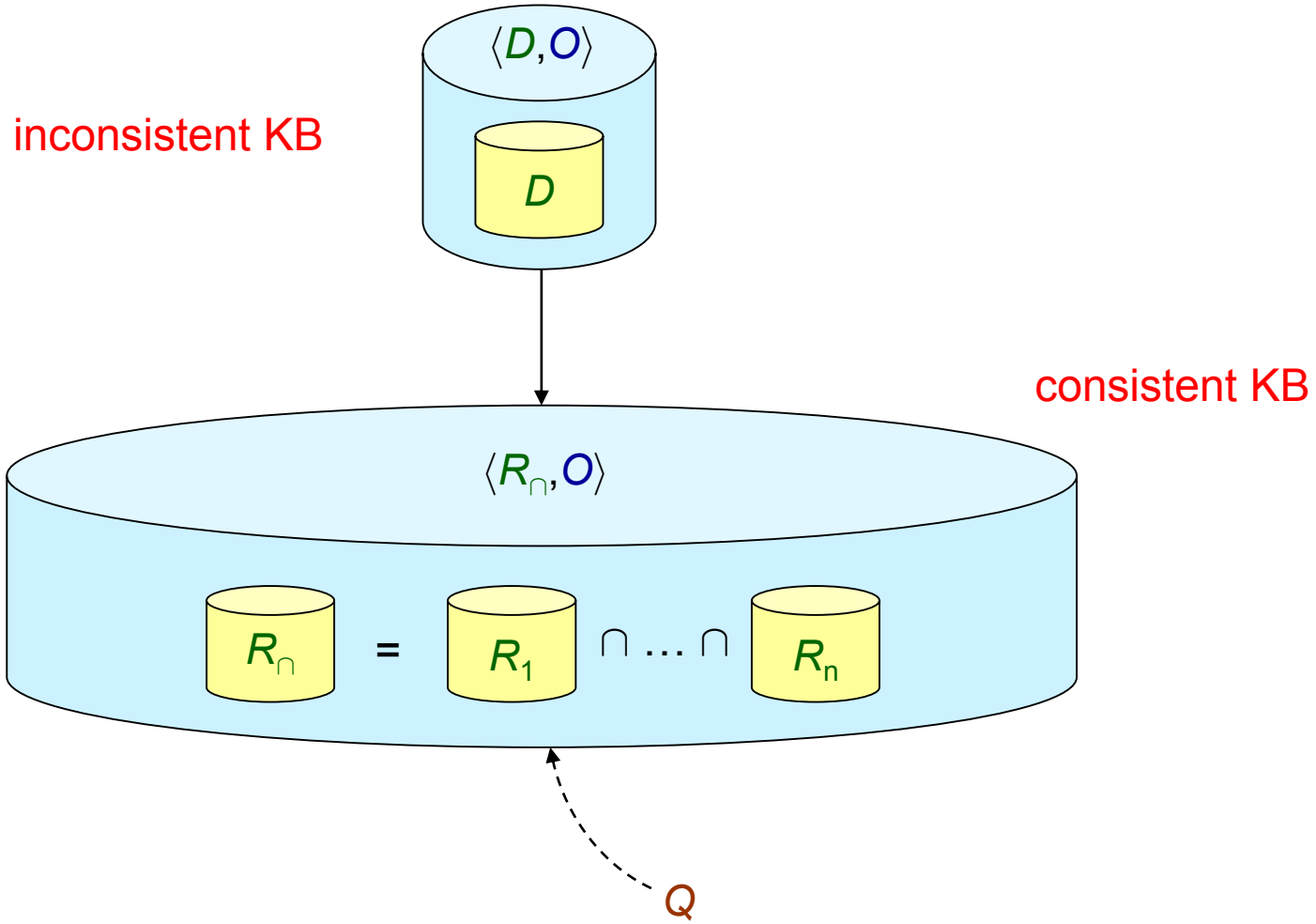
Intersection ABox Repair (IAR) Semantics

- One of the basic sound approximations of the AR semantics

- **IDEA:** The query must be entailed by the intersection of the database repairs

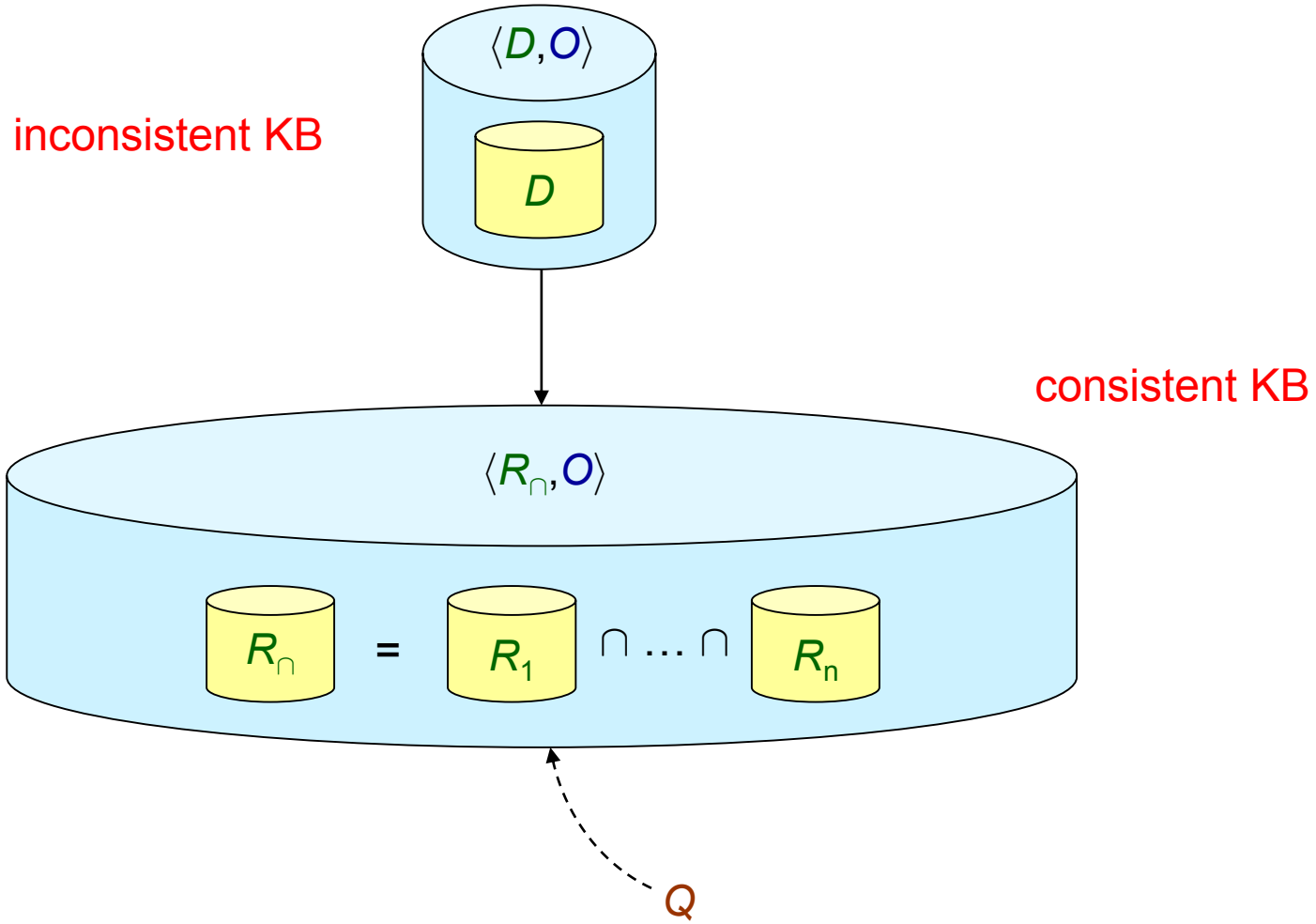
  ⊆-maximal consistent subsets of the database

Intersection ABox Repair (IAR) Semantics



IAR-answers(Q, $\langle D,\Sigma \rangle$) = certain-answers(Q, $\langle R_{\cap},\Sigma \rangle$)

Intersection ABox Repair (IAR) Semantics



IAR-answers(Q, $\langle D,\Sigma \rangle$) $\subseteq$ AR-answers(Q, $\langle D,\Sigma \rangle$)