Algebraic Aspects of Database Views: An Overview

Stephen J. Hegner
Umeå University
Department of Computing Science
Umeå, Sweden

http://www.cs.umu.se/~hegner

Motivation and Evolution

(Historical) motivation for the research:

Context: Research on the relational model 1970-1990

➪ particularly *decomposition theory*.

Question: To what extent are those results independent of the relational model, and applicable in a much wider context?

Evolution of the research program:

1983-1994: decomposition theory

1990, 2002-2005: updates and their complexity

2005-?: database components

Topic Overview

Part I: Schemata and views in the set based framework

- Develop the most basic ideas in an extremely general framework with a minimum of structure.

Part II: Schemata and views in structured frameworks

- Achieve results which are unobtainable in the pure set-based framework.
- The natural *order structure* which is present in most data models can be exploited to great advantage.

The Set-Based Framework

- First introduced by Bancilhon and Spyratos (1979-1981) in the study of independence and updates.

Schema: A *database schema* $\mathbf{D}$ is represented entirely by its set $\mathsf{LDB}(\mathbf{D})$ of *legal states*. All other information is "forgotten."

Morphism: A *database mapping* or *database morphism*

$$f : \mathbf{D}_1 \to \mathbf{D}_2$$

is just a function

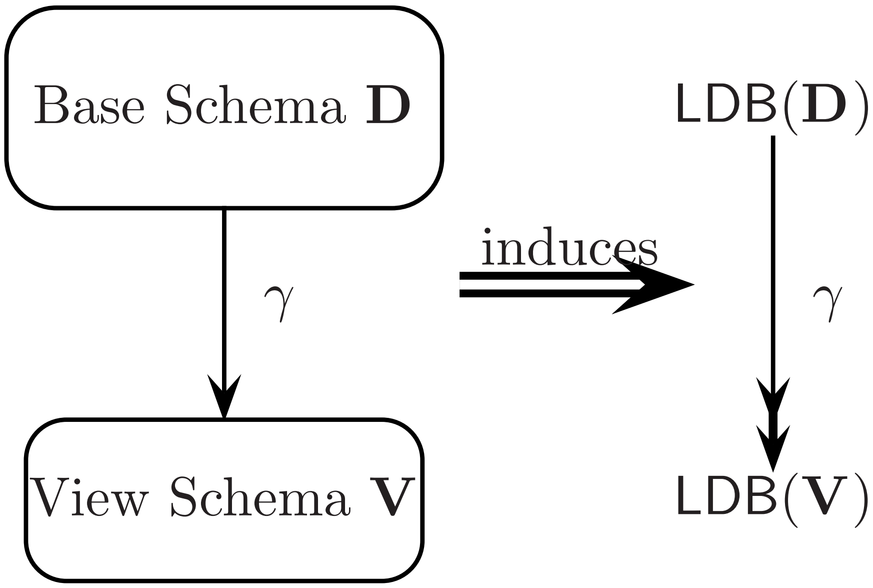
$$f : \mathsf{LDB}(\mathbf{D}_1) \to \mathsf{LDB}(\mathbf{D}_2)$$

- This recaptures (with some loss of structure) virtually all data models which do not involve object identity.

- A bit more on modelling object identifiers will come later.

Views in the Set-Based Framework

- A *view* of the schema $\mathbf{D}$ is a pair $\Gamma = (\mathbf{V}, \gamma)$ in which:
  - ➪ $\mathbf{V}$ is a schema;
  - ➪ $\gamma : \mathbf{D} \to \mathbf{V}$ is a *surjective* morphism.



View Morphisms

- For $\Gamma_1 = (\mathbf{V}_1, \gamma_1)$ and $\Gamma_2 = (\mathbf{V}_2, \gamma_2)$, a view *morphism*

$$f : \Gamma_1 \to \Gamma_2$$

  is a database morphism

$$f : \mathbf{V}_1 \to \mathbf{V}_2$$

  such that the following diagram commutes.

$$\begin{tikzcd} & \mathbf{D} \arrow[dl, "\gamma_1"'] \arrow[dr, "\gamma_2"] & \\ \mathbf{V}_1 \arrow[rr, "f"] & & \mathbf{V}_2 \end{tikzcd}$$

- A simple relational example:

$$\begin{tikzcd} & R[ABC] \arrow[dl, "\pi_{BC}"'] \arrow[dr, "\pi_C"] & \\ R[BC] \arrow[rr, "\pi_C"] & & R[C] \end{tikzcd}$$

Uniqueness of View Morphisms

Observation: There can be at most one morphism

$$f : \Gamma_1 \to \Gamma_2$$

Proof: Immediate from the surjectivity of γ_1. $\square$

- Morphisms can be composed.

$$\begin{tikzcd}
 & \mathbf{D} \arrow[dl, "\gamma_1"'] \arrow[d, "\gamma_2"] \arrow[dr, "\gamma_3"] & \\
\mathbf{V}_1 \arrow[r, "f"] \arrow[rr, bend right, dashed, "g \circ f"'] & \mathbf{V}_2 \arrow[r, "g"] & \mathbf{V}_3
\end{tikzcd}$$

- This induces a natural *preorder* (reflexive, transitive) on the set $\mathsf{View}(\mathbf{D})$ of all views of $\mathbf{D}$.

$$\Gamma_2 \leq \Gamma_1 \quad \text{iff} \quad (\exists f)(f : \Gamma_1 \to \Gamma_2)$$

- In turn, this induces a partial order on equivalence classes.

$$\Gamma_1 \equiv \Gamma_2 \quad \text{iff} \quad (\Gamma_1 \leq \Gamma_2 \text{ and } \Gamma_2 \leq \Gamma_1)$$
$$[\Gamma_1] \leq [\Gamma_2] \quad \text{iff} \quad \Gamma_1 \leq \Gamma_2$$

View Isomorphisms

Algebraic principle: Morphisms which connect objects which are "structurally identical" are called *isomorphisms*.

Definition: In general terms (not just for views), an *isomorphism* $f : A \to B$ has the property that there is another morphism $g : B \to A$ such that that both compositions are *identity morphisms*:

$$g \circ f = \mathbf{1}_A \quad \text{and} \quad f \circ g = \mathbf{1}_B$$

- In the context of sets and functions, isomorphisms are just bijective functions.
- This is also true for views, but more can be said.

Observation: If there are morphisms $f : \Gamma_1 \to \Gamma_2$ and $g : \Gamma_2 \to \Gamma_1$, then both are (view) isomorphisms. $\square$

$$\begin{tikzcd}
& & & \mathbf{D} \arrow[dlll, "\gamma_1"'] \arrow[dl, "\gamma_2"] \arrow[dr, "\gamma_1"'] \arrow[drrr, "\gamma_2"] & & & \\
\mathbf{V}_1 \arrow[rr, "f"] \arrow[rrrr, bend right, dashed, "\mathbf{1}_{\mathbf{V}_1}"'] & & \mathbf{V}_2 \arrow[rr, "g"] \arrow[rrrr, bend right, dashed, "\mathbf{1}_{\mathbf{V}_2}"'] & & \mathbf{V}_1 \arrow[rr, "f"] & & \mathbf{V}_2
\end{tikzcd}$$

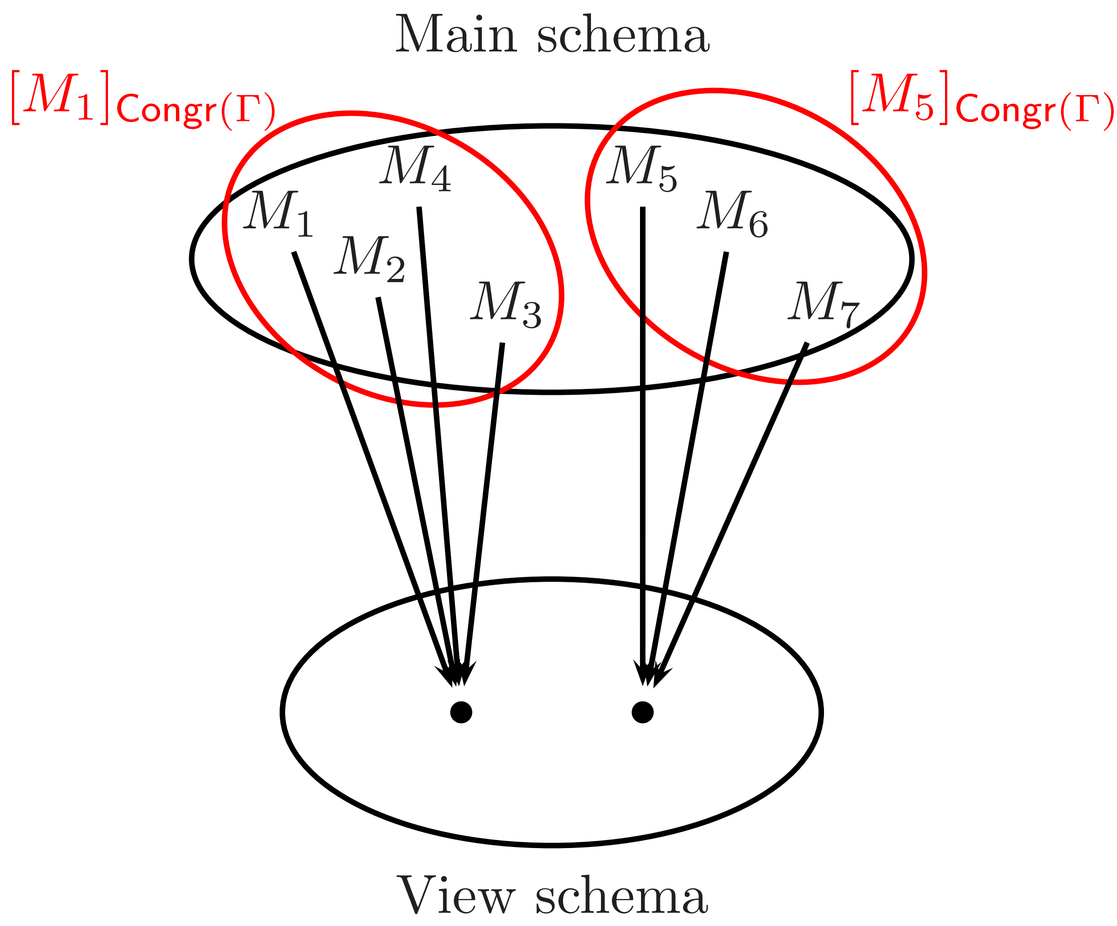
The Congruence of a View

Corollary: $[\Gamma_1]$ consists precisely of those views which are isomorphic to Γ_1. $\square$

Problem: Characterize $[\Gamma_1]$.

Definition: The *congruence* of $\Gamma = (\mathbf{V}, \gamma)$ is the (equivalence) relation

$$\mathsf{Congr}(\Gamma) =$$
$$\{(M_1, M_2) \in \mathsf{LDB}(\mathbf{D}) \times \mathsf{LDB}(\mathbf{D}) \mid \gamma(M_1) = \gamma(M_2)\}$$



The View Defined by a Congruence

- This idea may be "turned around" to obtain a view from an equivalence relation.

- Let $\mathbf{D} \;=\;$ schema;

  $R \;=\;$ any equivalence relation on $\mathsf{LDB}(\mathbf{D})$.

- Define $\Theta_R = (\mathbf{D}/R,\ \theta_R)$ with

  ➪ $\mathbf{D}/R$ = schema with $\mathsf{LDB}(\mathbf{D}/R) = \mathsf{LDB}(\mathbf{D})/R$

  = set of equivalence classes of R.

  ➪ $\theta_R : \mathbf{D} \to \mathbf{D}/R$ given by $M \mapsto [M]_R$.

Observation: Θ_R is a view of $\mathbf{D}$. $\square$

Observation: $[\Gamma] = [\Theta_{\mathsf{Congr}(\Gamma)}]$; *i.e.*, $\mathsf{Congr}(\Gamma)$ characterizes Γ, up to isomorphism. $\square$

$$\begin{tikzcd}
 & \mathbf{D} \arrow[dl, "\gamma"'] \arrow[dr, "\theta_{\mathsf{Congr}(\Gamma)}"] & \\
\mathbf{V} \arrow[rr, "\cong"] & & \mathbf{D}/\mathsf{Congr}(\Gamma)
\end{tikzcd}$$

$$\gamma(M) \longmapsto [M]_{\mathsf{Congr}(\Gamma)}$$

Congruences and Relative Views

<u>Key Observation</u>: There is a (unique) morphism

$$f : \Gamma_1 \to \Gamma_2$$

iff

$$\mathsf{Congr}(\Gamma_1) \subseteq \mathsf{Congr}(\Gamma_2). \qquad \square$$

$$\begin{tikzcd}
 & \mathbf{D} \arrow[dl, "\theta_{\mathsf{Congr}(\Gamma_1)}"'] \arrow[dr, "\theta_{\mathsf{Congr}(\Gamma_2)}"] & \\
\mathbf{D}/\mathsf{Congr}(\Gamma_1) \arrow[rr] & & \mathbf{D}/\mathsf{Congr}(\Gamma_2) \\
{[M]_{\mathsf{Congr}(\Gamma_1)}} \arrow[rr, mapsto, "\subseteq"] & & {[M]_{\mathsf{Congr}(\Gamma_2)}}
\end{tikzcd}$$

<u>Definition</u>: Given $\Gamma_1 = (\mathbf{V}_1, \gamma_1)$ and $\Gamma_2 = (\mathbf{V}_2, \gamma_2)$ with $\mathsf{Congr}(\Gamma_1) \subseteq \mathsf{Congr}(\Gamma_2)$, the *relative view*

$$\Lambda(\Gamma_1, \Gamma_2) = (\mathbf{V}_2, \lambda\langle\Gamma_1, \Gamma_2\rangle)$$

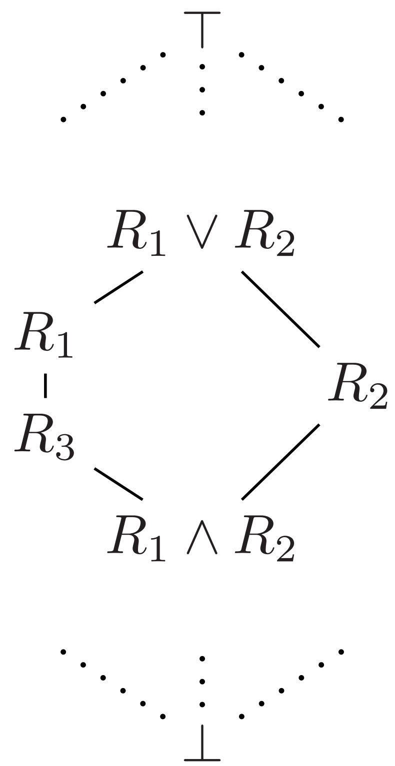
of $\mathbf{V}_1$ induced by Γ_2 is defined by the following diagram.

$$\begin{tikzcd}
 & \mathbf{D} \arrow[dl, "\gamma_1"'] \arrow[dr, "\gamma_2"] & \\
\mathbf{V}_1 \arrow[rr, "\lambda\langle\Gamma_1, \Gamma_2\rangle"] & & \mathbf{V}_2
\end{tikzcd}$$

The Lattice of Views of a Schema

Fact: (Øystein Ore 1942) The set of all equivalence relations $\mathcal{E}(X)$ on a set X forms a bounded lattice. $\square$

- $\top = \{(x, x) \mid x \in X\}$.
- $\bot = X \times X$.
- $R_1 \vee R_2 = R_1 \cap R_2$.
- $R_1 \wedge R_2 = \bigcap\{R \in \mathcal{E}(X) \mid R_1 \subseteq R \text{ and } R_2 \subseteq R\}$.



- This may also be regarded as a lattice of views over $\mathbf{D}$, with $X = \mathsf{LDB}(\mathbf{D})$.
  - ➪ $\top = \mathbf{1_D}$ = identity view.
  - ➪ $\bot = \mathbf{0_D}$ = zero view (The view schema has one state).

The Join Operation on Views

- The join operation subsumes and extends the classical operation of the relational theory.

Definition: Let $\Gamma_1 = (\mathbf{V}_1, \gamma_1)$ and $\Gamma_2 = (\mathbf{V}_2, \gamma_2)$ be views of $\mathbf{D}$. Define

$$\begin{aligned}
\Gamma_1 \times \Gamma_2 =& (\mathbf{V}_1 \,{}_{\gamma_1}\!\otimes_{\gamma_2} \mathbf{V}_2, \gamma_1 \otimes \gamma_2) \\
\mathsf{LDB}(\mathbf{V}_1 \,{}_{\gamma_1}\!\otimes_{\gamma_2} \mathbf{V}_2) =& \{(\gamma_1(M), \gamma_2(M)) \mid M \in \mathsf{LDB}(\mathbf{D})\} \\
\gamma_1 \otimes \gamma_2 : \mathbf{D} \to& \mathbf{V}_1 \,{}_{\gamma_1}\!\otimes_{\gamma_2} \mathbf{V}_2 \\
M \mapsto& (\gamma_1(M), \gamma_2(M))
\end{aligned}$$

- This construction is called a *subdirect product* in universal algebra.

Theorem: In the lattice of views of $\mathbf{D}$,

$$\Gamma_1 \vee \Gamma_2 = \Gamma_1 \times \Gamma_2 \qquad \square$$

- If $\Gamma_1 \times \Gamma_2 \cong \mathbf{1}_{\mathbf{D}}$, then
  - ➪ $\{\Gamma_1, \Gamma_2\}$ is called a *subdirect decomposition*.
  - ➪ $(\gamma_1 \otimes \gamma_2)^{-1} : \mathbf{V}_1 \,{}_{\gamma_1}\!\otimes_{\gamma_2} \mathbf{V}_2 \to \mathbf{D}$ is called the *reconstruction morphism*.

The Join Operation in the Relational Setting

- Example base schema: $(E[\mathsf{ABC}], \{\mathsf{A} \to \mathsf{C}, \mathsf{B} \to \mathsf{C}\})$

- Two views: $\Pi_{\mathsf{AB}} = ((E[\mathsf{AB}], \emptyset), \pi_{\mathsf{AB}})$
  $\Pi_{\mathsf{BC}} = ((E[\mathsf{BC}], \{\mathsf{B} \to \mathsf{C}\}), \pi_{\mathsf{BC}})$

- The decomposition mapping

$$\begin{aligned} \pi_{\mathsf{AB}} \otimes \pi_{\mathsf{BC}} : E[\mathsf{ABC}] &\to E[\mathsf{AB}]\ {}_{\pi_{\mathsf{AB}}}\!\otimes_{\pi_{\mathsf{BC}}} E[\mathsf{BC}] \\ M &\mapsto (\pi_{\mathsf{AB}}(M), \pi_{\mathsf{BC}}(M)) \end{aligned}$$

$$\begin{bmatrix} a_0 & b_0 & c_0 \\ a_1 & b_1 & c_1 \\ a_2 & b_1 & c_1 \end{bmatrix} \mapsto \left(\begin{bmatrix} a_0 & b_0 \\ a_1 & b_1 \\ a_2 & b_1 \end{bmatrix}, \begin{bmatrix} b_0 & c_0 \\ b_1 & c_1 \end{bmatrix} \right)$$

- The reconstruction mapping is just the (relational) join:

$$\begin{aligned} (\pi_{\mathsf{AB}} \otimes \pi_{\mathsf{BC}})^{-1} : E[\mathsf{AB}]\ {}_{\pi_{\mathsf{AB}}}\!\otimes_{\pi_{\mathsf{BC}}} E[\mathsf{BC}] &\to E[\mathsf{ABC}] \\ (\pi_{\mathsf{AB}}(M), \pi_{\mathsf{BC}}(M)) &\mapsto (\pi_{\mathsf{AB}}(M) \bowtie \pi_{\mathsf{BC}}(M)) \end{aligned}$$

$$\left(\begin{bmatrix} a_0 & b_0 \\ a_1 & b_1 \\ a_2 & b_1 \end{bmatrix}, \begin{bmatrix} b_0 & c_0 \\ b_1 & c_1 \end{bmatrix} \right) \mapsto \begin{bmatrix} a_0 & b_0 & c_0 \\ a_1 & b_1 & c_1 \\ a_2 & b_1 & c_1 \end{bmatrix}$$

The Meet Operation and Relative Views

Notation: For $\Gamma_1 = (\mathbf{V}_1, \gamma_1)$, $\Gamma_2 = (\mathbf{V}_2, \gamma_2)$.

$$\Gamma_1 \wedge \Gamma_2 = (\mathbf{V}_1 \; {}_{\gamma_1}\!\wedge_{\gamma_2} \mathbf{V}_2, \gamma_1 \wedge \gamma_2)$$

- Since

$$\begin{aligned} \Gamma_1 \wedge \Gamma_2 &\leq \Gamma_1 \\ \Gamma_1 \wedge \Gamma_2 &\leq \Gamma_2 \end{aligned}$$

  $\Gamma_1 \wedge \Gamma_2$ may be regarded as a view of both Γ_1 and Γ_2.

- The *relative views*

$$\begin{aligned} \Lambda(\Gamma_1, \Gamma_2) &= (\mathbf{V}_2, \lambda\langle\Gamma_1, \Gamma_2\rangle) \\ \Lambda(\Gamma_2, \Gamma_1) &= (\mathbf{V}_1, \lambda\langle\Gamma_1, \Gamma_2\rangle) \end{aligned}$$

  are characterized by the following diagram.

$$\begin{tikzcd}
& \mathbf{D} \arrow[dl, "\gamma_1"'] \arrow[dr, "\gamma_2"] \arrow[dd, "\gamma_1 \wedge \gamma_2"'] & \\
\mathbf{V}_1 \arrow[dr, "\lambda\langle\Gamma_1, \Gamma_2\rangle"'] & & \mathbf{V}_2 \arrow[dl, "\lambda\langle\Gamma_2, \Gamma_1\rangle"] \\
& \mathbf{V}_1 \; {}_{\gamma_1}\!\wedge_{\gamma_2} \mathbf{V}_2 &
\end{tikzcd}$$

View Meet in the Relational Context

Example base schema:

$(E[\mathsf{ABC}], \{\mathsf{A} \rightarrow \mathsf{C}, \mathsf{B} \rightarrow \mathsf{C}\})$:

A	B	C

View $\Pi_{\mathsf{AB}} = (E[\mathsf{AB}], \emptyset)$

A	B

View $\Pi_{\mathsf{BC}} = (E[\mathsf{BC}], \{\mathsf{B} \rightarrow \mathsf{C}\})$

B	C

$$
\begin{tikzcd}
& (E[\mathsf{ABC}], \{\mathsf{A} \rightarrow \mathsf{C}, \mathsf{B} \rightarrow \mathsf{C}\}) \arrow[dl, "\pi_{\mathsf{AB}}"'] \arrow[dd, "\pi_{\mathsf{AB}} \wedge \pi_{\mathsf{BC}} = \pi_{\mathsf{B}}"] \arrow[dr, "\pi_{\mathsf{BC}}"] & \\
(E[\mathsf{AB}], \emptyset) \arrow[dr, "\pi_{\mathsf{B}}"'] & & (E[\mathsf{BC}], \{\mathsf{B} \rightarrow \mathsf{C}\}) \arrow[dl, "\pi_{\mathsf{B}}"] \\
& (E[\mathsf{B}], \emptyset) &
\end{tikzcd}
$$

- The presences and/or absence of $\mathsf{A} \rightarrow \mathsf{C}$ determines further properties.

The Meet Operation and Commuting Congruences

Notation:

$$\begin{aligned}\mathsf{Congr}(\Gamma_1) \circ \mathsf{Congr}(\Gamma_2) = \{(M_1, M_2) \in \mathsf{LDB}(\mathbf{D}) \times \mathsf{LDB}(\mathbf{D})\\ \mid (\exists M_3 \in \mathsf{LDB}(\mathbf{D}))\\ (M_1, M_3) \in \mathsf{Congr}(\Gamma_1) \wedge (M_3, M_2) \in \mathsf{Congr}(\Gamma_2)\}\end{aligned}$$

- M_3 is a sort of "join" of $\gamma_1(M_1)$ and $\gamma_2(M_2)$.

Observation:

$$\mathsf{Congr}(\Gamma_1) \circ \mathsf{Congr}(\Gamma_2) \subseteq \mathsf{Congr}(\Gamma_1 \wedge \Gamma_2)$$
$$\mathsf{Congr}(\Gamma_2) \circ \mathsf{Congr}(\Gamma_1) \subseteq \mathsf{Congr}(\Gamma_1 \wedge \Gamma_2) \qquad \square$$

Definition: If

$$\mathsf{Congr}(\Gamma_1) \circ \mathsf{Congr}(\Gamma_2) = \mathsf{Congr}(\Gamma_2) \circ \mathsf{Congr}(\Gamma_1)$$

then $\{\Gamma_1, \Gamma_2\}$ is said to have *commuting congruences*, and is called a *fully commuting pair*.

Theorem: If (and only if) $\{\Gamma_1, \Gamma_2\}$ has commuting congruences, then

$$\mathsf{Congr}(\Gamma_1 \wedge \Gamma_2) = \mathsf{Congr}(\Gamma_1) \circ \mathsf{Congr}(\Gamma_2) \qquad \square$$

Example of Commuting Congruences

- Example base schema: $(E[\mathsf{ABC}], \{\mathsf{B} \to \mathsf{C}\})$:
- Two views: $\Pi_{\mathsf{AB}} = ((E[\mathsf{AB}], \emptyset), \pi_{\mathsf{AB}})$
  $\Pi_{\mathsf{BC}} = ((E[\mathsf{BC}], \{\mathsf{B} \to \mathsf{C}\}), \pi_{\mathsf{BC}})$

$$\left(\begin{bmatrix} a_0 & b_0 & c_0 \\ a_1 & b_1 & c_1 \end{bmatrix}, \begin{bmatrix} a_0 & b_0 & c_0 \\ a_1 & b_1 & c_0 \end{bmatrix} \right) \in \mathsf{Congr}(\Pi_{\mathsf{AB}})$$

$$\left(\begin{bmatrix} a_0 & b_0 & c_0 \\ a_1 & b_1 & c_0 \end{bmatrix}, \begin{bmatrix} a_0 & b_0 & c_0 \\ a_0 & b_1 & c_0 \end{bmatrix} \right) \in \mathsf{Congr}(\Pi_{\mathsf{BC}})$$

$$\Rightarrow$$

$$\left(\begin{bmatrix} a_0 & b_0 & c_0 \\ a_1 & b_1 & c_1 \end{bmatrix}, \begin{bmatrix} a_0 & b_0 & c_0 \\ a_0 & b_1 & c_0 \end{bmatrix} \right)$$

$$\in \mathsf{Congr}(\Pi_{\mathsf{AB}}) \circ \mathsf{Congr}(\Pi_{\mathsf{BC}}) \subseteq \mathsf{Congr}(\Pi_{\mathsf{AB}} \wedge \Pi_{\mathsf{BC}})$$

- Similarly

$$\left(\begin{bmatrix} a_0 & b_0 & c_0 \\ a_1 & b_1 & c_1 \end{bmatrix}, \begin{bmatrix} a_0 & b_0 & c_0 \\ a_0 & b_1 & c_0 \end{bmatrix} \right)$$

$$\in \mathsf{Congr}(\Pi_{\mathsf{BC}}) \circ \mathsf{Congr}(\Pi_{\mathsf{AB}})$$

- The following "join of projections" is in $\mathsf{LDB}((E[\mathsf{ABC}], \mathsf{B} \to \mathsf{C}))$.

$$\begin{bmatrix} a_0 & b_0 & c_0 \\ a_0 & b_1 & c_1 \end{bmatrix}$$

Example of Commuting Congruences

- Example base schema: $(E[\mathsf{ABC}], \{\mathsf{B} \to \mathsf{C}\})$:
- Two views: $\Pi_{\mathsf{AB}} = ((E[\mathsf{AB}], \emptyset), \pi_{\mathsf{AB}})$
  $\Pi_{\mathsf{BC}} = ((E[\mathsf{BC}], \{\mathsf{B} \to \mathsf{C}\}), \pi_{\mathsf{BC}})$

$$\left(\begin{bmatrix} a_0 & b_0 & c_0 \\ a_1 & b_1 & c_1 \end{bmatrix}, \begin{bmatrix} a_0 & b_0 & c_0 \\ a_1 & b_1 & c_0 \end{bmatrix} \right) \in \mathsf{Congr}(\Pi_{\mathsf{AB}})$$
$$\left(\begin{bmatrix} a_0 & b_0 & c_0 \\ a_1 & b_1 & c_0 \end{bmatrix}, \begin{bmatrix} a_0 & b_0 & c_0 \\ a_0 & b_1 & c_0 \end{bmatrix} \right) \in \mathsf{Congr}(\Pi_{\mathsf{BC}})$$
$$\Rightarrow$$
$$\left(\begin{bmatrix} a_0 & b_0 & c_0 \\ a_1 & b_1 & c_1 \end{bmatrix}, \begin{bmatrix} a_0 & b_0 & c_0 \\ a_0 & b_1 & c_0 \end{bmatrix} \right)$$
$$\in \mathsf{Congr}(\Pi_{\mathsf{AB}}) \circ \mathsf{Congr}(\Pi_{\mathsf{BC}}) \subseteq \mathsf{Congr}(\Pi_{\mathsf{AB}} \wedge \Pi_{\mathsf{BC}})$$

- Similarly

$$\left(\begin{bmatrix} a_0 & b_0 & c_0 \\ a_1 & b_1 & c_1 \end{bmatrix}, \begin{bmatrix} a_0 & b_0 & c_0 \\ a_0 & b_1 & c_0 \end{bmatrix} \right)$$
$$\in \mathsf{Congr}(\Pi_{\mathsf{BC}}) \circ \mathsf{Congr}(\Pi_{\mathsf{AB}})$$

- The following "join of projections" is in $\mathsf{LDB}((E[\mathsf{ABC}], \mathsf{B} \to \mathsf{C}))$.

$$\begin{bmatrix} a_0 & b_0 & c_0 \\ a_0 & b_1 & c_1 \end{bmatrix}$$

Example of Non-Commuting Congruences

- Example base schema: $(E[\mathsf{ABC}], \{\mathsf{A} \to \mathsf{C}, \mathsf{B} \to \mathsf{C}\})$:
- Two views: $\Pi_{\mathsf{AB}} = ((E[\mathsf{AB}], \emptyset), \pi_{\mathsf{AB}})$
  $\Pi_{\mathsf{BC}} = ((E[\mathsf{BC}], \{\mathsf{B} \to \mathsf{C}\}), \pi_{\mathsf{BC}})$

$$\left(\begin{bmatrix} a_0 & b_0 & c_0 \\ a_1 & b_1 & c_1 \end{bmatrix}, \begin{bmatrix} a_0 & b_0 & c_0 \\ a_1 & b_1 & c_0 \end{bmatrix} \right) \in \mathsf{Congr}(\Pi_{\mathsf{AB}})$$

$$\left(\begin{bmatrix} a_0 & b_0 & c_0 \\ a_1 & b_1 & c_0 \end{bmatrix}, \begin{bmatrix} a_0 & b_0 & c_0 \\ a_0 & b_1 & c_0 \end{bmatrix} \right) \in \mathsf{Congr}(\Pi_{\mathsf{BC}})$$

$$\Rightarrow$$

$$\left(\begin{bmatrix} a_0 & b_0 & c_0 \\ a_1 & b_1 & c_1 \end{bmatrix}, \begin{bmatrix} a_0 & b_0 & c_0 \\ a_0 & b_1 & c_0 \end{bmatrix} \right)$$

$$\in \mathsf{Congr}(\Pi_{\mathsf{AB}}) \circ \mathsf{Congr}(\Pi_{\mathsf{BC}}) \subseteq \mathsf{Congr}(\Pi_{\mathsf{AB}} \wedge \Pi_{\mathsf{BC}})$$

- However

$$\left(\begin{bmatrix} a_0 & b_0 & c_0 \\ a_1 & b_1 & c_1 \end{bmatrix}, \begin{bmatrix} a_0 & b_0 & c_0 \\ a_0 & b_1 & c_0 \end{bmatrix} \right)$$

$$\notin \mathsf{Congr}(\Pi_{\mathsf{BC}}) \circ \mathsf{Congr}(\Pi_{\mathsf{AB}})$$

- The following "join of projections" does not satisfy $\mathsf{A} \to \mathsf{C}$.

$$\begin{bmatrix} a_0 & b_0 & c_0 \\ a_0 & b_1 & c_1 \end{bmatrix}$$

Example of Non-Commuting Congruences

- Example base schema: $(E[\mathsf{ABC}], \{\mathsf{A} \to \mathsf{C}, \mathsf{B} \to \mathsf{C}\})$:
- Two views: $\Pi_{\mathsf{AB}} = ((E[\mathsf{AB}], \emptyset), \pi_{\mathsf{AB}})$
  $\Pi_{\mathsf{BC}} = ((E[\mathsf{BC}], \{\mathsf{B} \to \mathsf{C}\}), \pi_{\mathsf{BC}})$

$$\left(\begin{bmatrix} a_0 & b_0 & c_0 \\ a_1 & b_1 & c_1 \end{bmatrix}, \begin{bmatrix} a_0 & b_0 & c_0 \\ a_1 & b_1 & c_0 \end{bmatrix} \right) \in \mathsf{Congr}(\Pi_{\mathsf{AB}})$$

$$\left(\begin{bmatrix} a_0 & b_0 & c_0 \\ a_1 & b_1 & c_0 \end{bmatrix}, \begin{bmatrix} a_0 & b_0 & c_0 \\ a_0 & b_1 & c_0 \end{bmatrix} \right) \in \mathsf{Congr}(\Pi_{\mathsf{BC}})$$

$$\Rightarrow$$

$$\left(\begin{bmatrix} a_0 & b_0 & c_0 \\ a_1 & b_1 & c_1 \end{bmatrix}, \begin{bmatrix} a_0 & b_0 & c_0 \\ a_0 & b_1 & c_0 \end{bmatrix} \right)$$

$$\in \mathsf{Congr}(\Pi_{\mathsf{AB}}) \circ \mathsf{Congr}(\Pi_{\mathsf{BC}}) \subseteq \mathsf{Congr}(\Pi_{\mathsf{AB}} \wedge \Pi_{\mathsf{BC}})$$

- However

$$\left(\begin{bmatrix} a_0 & b_0 & c_0 \\ a_1 & b_1 & c_1 \end{bmatrix}, \begin{bmatrix} a_0 & b_0 & c_0 \\ a_0 & b_1 & c_0 \end{bmatrix} \right)$$

$$\notin \mathsf{Congr}(\Pi_{\mathsf{BC}}) \circ \mathsf{Congr}(\Pi_{\mathsf{AB}})$$

- The following "join of projections" does not satisfy $\mathsf{A} \to \mathsf{C}$.

$$\begin{bmatrix} a_0 & b_0 & c_0 \\ a_0 & b_1 & c_1 \end{bmatrix}$$

Reconstruction and Independence Dependencies

Context: $\Gamma_1 = (\mathbf{V}_1, \gamma_1)$ and $\Gamma_2 = (\mathbf{V}_2, \gamma_2)$ views of $\mathbf{D}$.

- The $\{\Gamma_1, \Gamma_2\}$-*reconstruction dependency* on $\mathbf{V}_1 \,{}_{\gamma_1}\!\!\otimes_{\gamma_2} \mathbf{V}_2$:
  - ➪ Denoted $\otimes[\Gamma_1, \Gamma_2]$.
  - ➪ Satisfied if whenever:
    - $M_1, N_1 \in \mathsf{LDB}(\mathbf{V}_1)$ and $M_2, N_2 \in \mathsf{LDB}(\mathbf{V}_2)$,
  - ➪ then if any three of the elements of the set $\{(M_1, M_2), (M_1, N_2), (N_1, M_2), (N_1, N_2)\}$ are in $\mathsf{LDB}(\mathbf{V}_1 \,{}_{\gamma_1}\!\!\otimes_{\gamma_2} \mathbf{V}_2)$, so too is the fourth.
- The $(\Gamma_1 \wedge \Gamma_2)$-*independence dependency* on $\mathbf{V}_1 \,{}_{\gamma_1}\!\!\otimes_{\gamma_2} \mathbf{V}_2$:
  - ➪ Denoted $\otimes_{\Gamma_1 \wedge \Gamma_2}$.
  - ➪ Satisfied iff for any

    $M_1 \in \mathsf{LDB}(\mathbf{V}_1)$ and $M_2 \in \mathsf{LDB}(\mathbf{V}_2)$:

$$((M_1, M_2) \in \mathsf{LDB}(\mathbf{V}_1 \,{}_{\gamma_1}\!\!\otimes_{\gamma_2} \mathbf{V}_2)) \Leftrightarrow$$
$$(\lambda\langle \Gamma_1 1, \Gamma_1 \wedge \Gamma_2\rangle(M_1) = \lambda\langle \Gamma_2, \Gamma_1 \wedge \Gamma_2\rangle(M_2))$$

Theorem: The following are equivalent:

(a) $\{\Gamma_1, \Gamma_2\}$ is a fully commuting pair.

(b) $\mathbf{V}_1 \,{}_{\gamma_1}\!\!\otimes_{\gamma_2} \mathbf{V}_2$ satisfies $\otimes[\Gamma_1, \Gamma_2]$.

(c) $\mathbf{V}_1 \,{}_{\gamma_1}\!\!\otimes_{\gamma_2} \mathbf{V}_2$ satisfies $\otimes_{\Gamma_1 \wedge \Gamma_2}$. □

Reconstruction, Independence, and Dependencies

Theorem: In the relational setting, for a wide variety of views, $\{\Gamma_1, \Gamma_2\}$ is a fully commuting pair iff the decomposition is both lossless and dependency preserving. $\square$

Meta-conjecture: This result extends to any framework in which a reasonable rendering of *dependency preserving* may be established.

A Partial Lattice of Views

- Here is some terminology and notation for the curious.

Definition: A (bounded) *partial lattice* is similar to a lattice, save that the operations may be partial.

  - A true partial lattice may be extended to a lattice.
  - A *weak partial lattice* need not be so extendable.
  - See G. Grätzer, *General Lattice Theory*, for more details.

Definition: Define

$$\Gamma_1 \curlywedge \Gamma_2 = \begin{cases} \Gamma_1 \wedge \Gamma_2 & \text{if } \{\Gamma_1, \Gamma_2\} \text{ fully commuting pair.} \\ \text{undefined} & \text{otherwise.} \end{cases}$$

- The set of all views of a schema $\mathbf{D}$ may be regarded as a bounded partial lattice with the same join as defined earlier, but with meet as defined above.

Acyclic Decompositions

- This application is formulated entirely in the context of set-based views.

Example: $\mathbf{E} = (R[\mathsf{ABCD}], \{\mathsf{A} \rightarrow \mathsf{C}, \mathsf{B} \rightarrow \mathsf{D}\})$

Views in decomposition: Π_{AB}, Π_{ACD}, Π_{BCD}.

Observation: All of the attributes in ABCD are contained in any join of two components:

$$\mathsf{AB} \bowtie \mathsf{ACD}, \quad \mathsf{AB} \bowtie \mathsf{BCD}, \quad \mathsf{ACD} \bowtie \mathsf{BCD}$$

Fact: None of these binary joins recovers the entire relation; *i.e.*, none is lossless.

Fact: The join $\bowtie [\mathsf{AB}, \mathsf{ACD}, \mathsf{BCD}])$ is lossless.

- This type of decomposition was studied extensively in the 1980's using the underlying *hypergraph* of the schema as a base.
- Such decompositions have *cyclic* hypergraphs, while "good" decompositions have *acyclic* hypergraphs.

Acyclic Decompositions Continued

- In (Hegner 1991 MFDBS) and (Hegner 1993 AMAI), it is shown that these results are not dependent in any way upon the relational model.
- There are 28 related characterizations in this theory.
- All are related to the properties of the view congruences, and nothing more.
- One interesting characterization is based upon a generalization of the famous *Chinese Remainder Theorem* of number theory.

Decomposition and Complements – The Need for More Structure

Definition: The pair $\{\Gamma_1, \Gamma_2\}$ of views forms a *(subdirect) complementary pair* if $\Gamma_1 \vee \Gamma_2 = \mathbf{1_D}$.

- In this case, Γ_2 is called a *(subdirect) complement* of Γ_1, and vice-versa.
- If $\{\Gamma_1, \Gamma_2\}$ is furthermore fully commuting, then it is called a *meet-complementary pair*.

Problem: Complements are almost never unique in the pure set-based framework.

- This is true even though there is often a "natural" and intuitive complement.
- The solution is to make use of "natural" additional structure of the schemata and morphisms; typically order structure.
- First, a pair of examples is given to illustrate the point.

Example 1: Order-Respecting Views

Base schema:

$(R[\mathsf{A}], \emptyset)$:

A
a_0
a_1
a_2

$(S[\mathsf{A}], \emptyset)$:

A
a_1
a_3
a_4

A View:

View $\Pi_{R[\mathsf{A}]}$:

A
a_0
a_1
a_2

Natural Complement:

View $\Pi_{S[\mathsf{A}]}$:

A
a_1
a_3
a_4

Alternative Complement:

View $\Pi_{R\Delta S[\mathsf{A}]}$:

A
a_0
a_2
a_3
a_4

- $R\Delta S$ is the *symmetric difference* of R and S.

$$R\Delta S[\mathsf{A}] = (R[\mathsf{A}] \cup S[\mathsf{A}]) \setminus (R[\mathsf{A}] \cap S[\mathsf{A}])$$

Question: How may $\Pi_{R[\mathsf{A}]}$ be favored over $\Pi_{R\Delta S[\mathsf{A}]}$ on formal grounds?

Rough answer: $\Pi_{S[\mathsf{A}]}$ is isotonic with respect to the natural order structure, while $\Pi_{R\Delta S[\mathsf{A}]}$ is not.

Example 2: Order-Respecting Views

Base schema:

$(E[\mathsf{ABC}], \{\mathsf{B} \to \mathsf{C}\})$

A	B	C
a_0	b_0	c_0
a_1	b_1	c_1
a_2	b_1	c_1

A View:

View $\Pi_{\mathsf{AB}} = (E[\mathsf{AB}], \emptyset)$

A	B
a_0	b_0
a_1	b_1
a_2	b_1

- For any $b \in \mathsf{Dom}(\mathsf{B})$, $\#_{\mathsf{A}}(b)$ = number of distinct values of for attribute A associated with b in $\pi_{\mathsf{AB}}(M)$.
- Let $\mathsf{Dom}(C) = \{c_0, c_1, c_2\}$.
- $\alpha : \mathsf{Dom}(C) \to \mathsf{Dom}(C)$; $c_i \mapsto c_{(i+1 \mod 3)}$.
- Define $\pi'_{BC}(M) =$
  $\{(b, c) \mid (b, c) \in \pi_{BC}(M) \text{ and } \#_A(b) \text{ is odd}\} \cup$
  $\{(b, \alpha(c)) \mid (b, c) \in \pi_{BC}(M) \text{ and } \#_A(b) \text{ is even}\}$.
- Π'_{BC} is a rather complex but perfectly correct meet complement to Π_{AB}.
- Unlike Π_{BC}, it does not respect order.

Complementary view:

$\Pi'_{\mathsf{BC}} = (E[\mathsf{BC}], \{\mathsf{B} \to \mathsf{C}\})$

B	C
b_0	c_0
b_1	c_2

Schemata in an Order Framework

Definition: An *order database schema* is just a poset:

$$\mathbf{D} = (\mathsf{LDB}(\mathbf{D}), \leq_{\mathbf{D}})$$

- There are many variations on the basic paradigm.

➪ A *⊥-schema* (Hegner JCSS 1994) is a triple

$$\mathbf{D} = (\mathsf{LDB}(\mathbf{D}), \leq_{\mathbf{D}}, \bot_{\mathbf{D}})$$

in which $\bot_{\mathbf{D}}$ is the least element in the ordering.

➪ A *PF-schema* (Hegner AMAI 2005) is a pair

$$\mathbf{D} = (\mathsf{SynFnd}(\mathbf{D}), \mathsf{LDB}(\mathbf{D}))$$

in which $\mathsf{SynFnd}(\mathbf{D})$ is a set and $\mathsf{LDB}(\mathbf{D})$ is a subset of the set of all finite subsets of $\mathsf{SynFnd}(\mathbf{D})$.

- Think of $\mathsf{SynFnd}(\mathbf{D})$ as a set of "tuples" or "entities" or "objects".
- Think of $\mathsf{LDB}(\mathbf{D})$ as a set of finite relations over $\mathsf{SynFnd}(\mathbf{D})$.
- $\mathsf{LDB}(\mathbf{D})$ has the natural ordering from set inclusion $\subseteq$.

Morphisms in an Order Framework

Definition: An *order database morphism*

$$f : \mathbf{D}_1 \to \mathbf{D}_2$$

is just a poset morphism

$$f : \mathsf{LDB}(\mathbf{D}_1) \to \mathsf{LDB}(\mathbf{D}_2)$$

- There are many variations on the basic paradigm.

➪ A *⊥-schema morphism* preserves the least element.

$$f(\bot_{\mathbf{D}_1}) = \bot_{\mathbf{D}_2}$$

➪ A *PF-morphism* is more complex. The idea that the (order) mapping

$$f^{\flat} : \mathsf{LDB}(\mathbf{D}_1) \to \mathsf{LDB}(\mathbf{D}_2)$$

is induced on a "tuple-by-tuple" basis from an associated mapping

$$f^{\sharp} : \mathsf{SynFnd}(\mathbf{D}_1) \to \mathsf{SynFnd}(\mathbf{D}_2)$$

  - This recaptures projection and selection from the relational setting, but not join.

Views and Congruences in an Order Framework

- The definition of *view* is entirely analogous.
- The definition of congruence must be specialized to an appropriate *order congruence*.
- With the specialization, the following still hold.

Proposition: There is a (unique) morphism $f : \Gamma_1 \to \Gamma_2$ iff $\mathsf{Congr}(\Gamma_1) \subseteq \mathsf{Congr}(\Gamma_2)$; *i.e.*, the dashed fill-in is an order morphism. $\square$

$$\begin{tikzcd}
 & \mathbf{D} \arrow[dl, "\theta_{\mathsf{Congr}(\Gamma_1)}"'] \arrow[dr, "\theta_{\mathsf{Congr}(\Gamma_2)}"] & \\
\mathbf{D}/\mathsf{Congr}(\Gamma_1) \arrow[rr, dashed] & & \mathbf{D}/\mathsf{Congr}(\Gamma_2) \\
{[M]_{\mathsf{Congr}(\Gamma_1)}} \arrow[rr, mapsto, "\subseteq"] & & {[M]_{\mathsf{Congr}(\Gamma_2)}}
\end{tikzcd}$$

➪ Relative views may be defined as in the set-based case.

➪ There is a natural bounded lattice structure on the views.

Identity-Preserving Views

- In general, views defined as in these lectures do not preserve object identity.
- However, it is possible to define a (possibly useful?) subclass of PF-views which does preserve such identity.
- To accomplish this, it is necessary to employ a model of database schemata which recaptures tuples, entities, and/or objects, as well as databases themselves.
- Recall:

➪ A *PF-schema* (Hegner AMAI 2005) is a pair

$$\mathbf{D} = (\mathsf{SynFnd}(\mathbf{D}), \mathsf{LDB}(\mathbf{D}))$$

in which $\mathsf{SynFnd}(\mathbf{D})$ is a set and $\mathsf{LDB}(\mathbf{D})$ is a subset of the set of all finite subsets of $\mathsf{SynFnd}(\mathbf{D})$.

- Think of $\mathsf{SynFnd}(\mathbf{D})$ as a set of "tuples" or "entities" or "objects".
- Think of $\mathsf{LDB}(\mathbf{D})$ as a set of finite relations over $\mathsf{SynFnd}(\mathbf{D})$.
- $\mathsf{LDB}(\mathbf{D})$ has the natural ordering from set inclusion $\subseteq$.

Example of a PF-schema

Main schema: Let $\mathbf{R}$ have two relations: $R_1[\mathsf{ABC}]$, $R_2[\mathsf{CDE}]$, with $\mathsf{A} \to \mathsf{BC}$ (on $R_1[\mathsf{ABC}]$).

- The (possibly infinite) domains of A, B, C, D, and E are $\mathsf{Dom}(\mathsf{A})$, $\mathsf{Dom}(\mathsf{B})$, $\mathsf{Dom}(\mathsf{C})$, $\mathsf{Dom}(\mathsf{E})$, and $\mathsf{Dom}(\mathsf{F})$.
- $$\begin{array}{ll} & \mathsf{SynFnd}(\mathbf{R}) = \\ & \{{}_1(a,b,c) \mid a \in \mathsf{Dom}(\mathsf{A}) \wedge b \in \mathsf{Dom}(\mathsf{B}) \wedge c \in \mathsf{Dom}(\mathsf{C})\ \} \\ \cup & \{{}_2(c,d,e) \mid c \in \mathsf{Dom}(\mathsf{C}) \wedge d \in \mathsf{Dom}(\mathsf{D}) \wedge e \in \mathsf{Dom}(\mathsf{E})\ \} \end{array}$$

View schema: $\mathbf{S}$ has just one relation $S[\mathsf{AB}]$, $\mathsf{A} \to \mathsf{B}$.

- $\mathsf{SynFnd}(\mathbf{S}) = \{(a,b) \mid a \in \mathsf{Dom}(\mathsf{A}) \wedge b \in \mathsf{Dom}(\mathsf{B})\}$

Null tuples:
$$\begin{array}{ll} \mathsf{ExtSynFnd}(\mathbf{R}) & = \mathsf{SynFnd}(\mathbf{R}) \cup \{\eta_{\mathbf{R}}\} \\ \mathsf{ExtSynFnd}(\mathbf{S}) & = \mathsf{SynFnd}(\mathbf{S}) \cup \{\eta_{\mathbf{S}}\} \end{array}$$

View morphism on elements:

$$\begin{array}{rcccc} \omega^\sharp & : & \mathsf{ExtSynFnd}(\mathbf{R}) & \to & \mathsf{ExtSynFnd}(\mathbf{S}) \\ & & {}_1(a,b,c) & \mapsto & (a,b) \\ & & {}_2(c,d,e) & \mapsto & \eta_{\mathbf{S}} \\ & & \eta_{\mathbf{R}} & \mapsto & \eta_{\mathbf{S}} \end{array}$$

View morphism on (finite) databases:

$$\omega^\flat : \mathsf{LDB}(\mathbf{R}) \to \mathsf{LDB}(\mathbf{S})$$

is just the projection of $R_1[\mathsf{ABC}]$ onto $S[\mathsf{AB}]$.

➪ $R_2[\mathsf{CDE}]$ is ignored.

Example: Identity-Preserving Views

Observation: For each pair $t_1, t_2 \in \mathsf{ExtSynFnd}(\mathbf{R})$, at least one of of the following three properties holds:

(i) $\omega^\sharp(t_1) \neq \omega^\sharp(t_2)$.

(ii) $\omega^\sharp(t_1) = \eta_{\mathrm{S}}$

(iii) $\omega^\sharp(t_2) = \eta_{\mathrm{S}}$.

- Informally, tuples which are not discarded have their identity preserved.
- This leads to an obvious general definition of an *identity-preserving view*.

➪ The view schema consists of some of the objects of the main schema, with perhaps some attributes (not affecting identity) discarded.

Question: Is this concept useful?

Uniqueness of Complements

- Order database schemata and views in general, and PF-schemata and PF-views in particular, are very useful in forcing (meet) complements to be unique.
- This property is very important in modelling view update strategies.
- The details will be presented in a separate talk.

➪ Unique complements are also important in database decomposition.

  - A decomposition is modelled as a Boolean subalgebra of the partial lattice of all views with commuting congruences.
  - See [Hegner JCSS 1994] for details.