

# Relational Division in Rank-Aware Databases: An Overview, Issues, and New Directions

Ondrej Vaverka, Vilem Vychodil

Dept. Computer Science, Palacky University, Olomouc, Czech Republic

## Rank-Aware Databases

- ▶ Rank-aware databases allow imperfect matches.
- ▶ Each tuple is annotated by a score – a number quantifying how much a tuple matches the query.

### Example

| <i>id</i> | <i>car</i>     | <i>color</i> | <i>price</i> |
|-----------|----------------|--------------|--------------|
| 1         | Honda Civic    | blue         | \$18.500     |
| 2         | Mazda 323F     | dark red     | \$17.500     |
| 3         | Toyota Celica  | light blue   | \$21.000     |
| 4         | Toyota Corolla | blue         | \$20.000     |

**Intended query:** Get all blue cars with price around \$20.000.

## Rank-Aware Databases

- ▶ Rank-aware databases allow imperfect matches.
- ▶ Each tuple is annotated by a score – a number quantifying how much a tuple matches the query.

### Example

| <i>id</i> | <i>car</i>     | <i>color</i> | <i>price</i> |
|-----------|----------------|--------------|--------------|
| 1         | Honda Civic    | blue         | \$18.500     |
| 2         | Mazda 323F     | dark red     | \$17.500     |
| 3         | Toyota Celica  | light blue   | \$21.000     |
| 4         | Toyota Corolla | blue         | \$20.000     |

**Query:** Get all blue cars with price between \$19.000 and \$21.000.

## Rank-Aware Databases

- ▶ Rank-aware databases allow imperfect matches.
- ▶ Each tuple is annotated by a score – a number quantifying how much a tuple matches the query.

### Example

| <i>score</i> | <i>id</i> | <i>car</i>     | <i>color</i> | <i>price</i> |
|--------------|-----------|----------------|--------------|--------------|
| 1.00         | 4         | Toyota Corolla | blue         | \$20.000     |
| 0.85         | 1         | Honda Civic    | blue         | \$18.500     |
| 0.80         | 3         | Toyota Celica  | light blue   | \$21.000     |
| 0.15         | 2         | Mazda 323F     | dark red     | \$17.500     |

**Query:** Get all blue cars with price around \$20.000.

**Note:** We call such table a "ranked data table" (RDT).

## Approach Based on Residuated Lattices

- ▶ There are many approaches to rank-aware databases that differ in the treatment of the scores of tuples.
- ▶ In our model the tuple scores come from residuated lattices.
- ▶ Our model can be seen as a generalization of Codd's relational model of data.

### Definition

Residuated lattice is a general algebra of the form

$$\mathbf{L} = \langle L, \wedge, \vee, \otimes, \rightarrow, 0, 1 \rangle$$

where

- ▶  $\langle L, \wedge, \vee, 0, 1 \rangle$  is a bounded lattice,
- ▶  $\langle L, \otimes, 1 \rangle$  is a commutative monoid,
- ▶ multiplication  $\otimes$  and residuum  $\rightarrow$  satisfy adjointness property.

## Query Systems

Two basic types of query systems:

1. system based on evaluating predicate formulas,
2. system consisting of relational operations.

For convenience:

Abbreviation  $RS$  denotes relational schema  $R \cup S$  with  $R \cap S = \emptyset$ .

## Query Systems

Two basic types of query systems:

1. system based on evaluating predicate formulas,
2. system consisting of relational operations.

For convenience:

Abbreviation  $RS$  denotes relational schema  $R \cup S$  with  $R \cap S = \emptyset$ .

### Examples of relational operations in our model

- ▶ Natural Join of RDTs  $\mathcal{D}_1$  (on  $RS$ ) and  $\mathcal{D}_2$  (on  $ST$ ) is an RDT on  $RST$  given by

$$(\mathcal{D}_1 \bowtie \mathcal{D}_2)(rst) = \mathcal{D}_1(rs) \otimes \mathcal{D}_2(st)$$

for all  $r \in \text{Tupl}(R)$ ,  $s \in \text{Tupl}(S)$ ,  $t \in \text{Tupl}(T)$ .

## Query Systems

Two basic types of query systems:

1. system based on evaluating predicate formulas,
2. system consisting of relational operations.

For convenience:

Abbreviation  $RS$  denotes relational schema  $R \cup S$  with  $R \cap S = \emptyset$ .

### Examples of relational operations in our model

- ▶ Natural Join of RDTs  $\mathcal{D}_1$  (on  $RS$ ) and  $\mathcal{D}_2$  (on  $ST$ ) is an RDT on  $RST$  given by

$$(\mathcal{D}_1 \bowtie \mathcal{D}_2)(rst) = \mathcal{D}_1(rs) \otimes \mathcal{D}_2(st)$$

for all  $r \in \text{Tupl}(R), s \in \text{Tupl}(S), t \in \text{Tupl}(T)$ .

- ▶ Projection of RDT  $\mathcal{D}$  (on  $R$ ) onto  $S \subseteq R$  is defined by

$$(\pi_S(\mathcal{D}))(s) = \bigvee_{t \in \text{Tupl}(R \setminus S)} \mathcal{D}(st)$$

for all  $s \in \text{Tupl}(S)$ .



## "Some $\varphi$ Is $\psi$ " Query

- ▶ Query: Some tuples from  $\mathcal{D}_1$  are matching tuples in  $\mathcal{D}_2$ .
- ▶ Such queries can be expressed using joins and projections.

### Example

Employees

| <i>name</i> | ... |
|-------------|-----|
| Carolyn     | ... |
| Amy         | ... |
| Jean        | ... |
| Elizabeth   | ... |

Teaching assignments

| <i>name</i> | <i>course</i>    |
|-------------|------------------|
| Amy         | Database systems |
| Amy         | Algorithms       |
| Jean        | Compilers        |
| Peter       | Machine Learning |

**Query:** Employees that teach some courses.

- ▶  $\pi_{\{name, \dots\}}(\text{Employees} \bowtie \text{Teaching assignments})$
- ▶  $\text{Employees} \bowtie \pi_{\{name\}}(\text{Teaching assignments})$

## "All $\varphi$ Are $\psi$ " Query

- ▶ In the classical model this type of query is connected to relational division.

### Example

Supplies

| <i>supplier</i> | <i>product</i> |
|-----------------|----------------|
| Logitech        | Keyboard       |
| Logitech        | Microphone     |
| Logitech        | Webcam         |
| IBM             | Keyboard       |
| IBM             | Webcam         |

Required Products

| <i>product</i> |
|----------------|
| Keyboard       |
| Microphone     |
| Webcam         |

**Query:** Suppliers that supply all required products.

- ▶  $\text{Supplies} \div \text{Required Products}$

## "All $\varphi$ Are $\psi$ " Query

- ▶ In the classical model this type of query is connected to relational division.

### Example

Supplies

| <i>supplier</i> | <i>product</i> |
|-----------------|----------------|
| Logitech        | Keyboard       |
| Logitech        | Microphone     |
| Logitech        | Webcam         |
| IBM             | Keyboard       |
| IBM             | Webcam         |

Required Products

| <i>product</i> |
|----------------|
| Keyboard       |
| Microphone     |
| Webcam         |

**Query:** Suppliers that supply all required products.

- ▶  $\text{Supplies} \div \text{Required Products} =$ 

|                 |
|-----------------|
| <i>supplier</i> |
| Logitech        |

## Codd-style Division

- ▶ The Codd-style division is an initial operation in the class of division-like operations.
- ▶ Its presence in the classical relational algebra is not necessary due to the fact that  $(\forall x)\varphi \equiv \neg(\exists x)\neg\varphi$ .
- ▶ In the classical model we consider division as a derived operation expressed by the means of set difference and projection.

### Definition

For relation  $\mathcal{D}_1$  on  $RS$  and relation  $\mathcal{D}_2$  on  $S$ , the Codd-style division may be introduced as

$$\mathcal{D}_1 \div_{\text{Codd}} \mathcal{D}_2 = \pi_R(\mathcal{D}_1) \setminus \pi_R((\pi_R(\mathcal{D}_1) \bowtie \mathcal{D}_2) \setminus \mathcal{D}_1).$$

## Issues of Codd-style Division

1. It is restricted to relations on particular schemes which makes the operation less general.

## Issues of Codd-style Division

1. It is restricted to relations on particular schemes which makes the operation less general.
2. The meaning of the operation does not faithfully correspond to the categorical proposition "all  $\varphi$  are  $\psi$ ".

## Issues of Codd-style Division

1. It is restricted to relations on particular schemes which makes the operation less general.
2. The meaning of the operation does not faithfully correspond to the categorical proposition "all  $\varphi$  are  $\psi$ ".

If  $\varphi$  is  $s \in \mathcal{D}_2$  and  $\psi$  is  $rs \in \mathcal{D}_1$ , then

$$(\forall s)(s \in \mathcal{D}_2 \Rightarrow rs \in \mathcal{D}_1)$$

is true for all  $r \in \text{Tupl}(R)$  if  $\mathcal{D}_2$  is empty.

## Issues of Codd-style Division

1. It is restricted to relations on particular schemes which makes the operation less general.
2. The meaning of the operation does not faithfully correspond to the categorical proposition "all  $\varphi$  are  $\psi$ ".

If  $\varphi$  is  $s \in \mathcal{D}_2$  and  $\psi$  is  $rs \in \mathcal{D}_1$ , then

$$(\forall s)(s \in \mathcal{D}_2 \Rightarrow rs \in \mathcal{D}_1)$$

is true for all  $r \in \text{Tuple}(R)$  if  $\mathcal{D}_2$  is empty.

The result of Codd-style division is always a subset of  $\pi_R(\mathcal{D}_1)$  and can be characterized as

$$\mathcal{D}_1 \div_{\text{Codd}} \mathcal{D}_2 = \{r \in \pi_R(\mathcal{D}_1) \mid \text{for all } s \in \mathcal{D}_2, \text{ we have } rs \in \mathcal{D}_1\}.$$



## Division in Rank-Aware Databases

- ▶ Most common approaches to rank-aware databases found in literature introduce the division by

$$(\mathcal{D}_1 \div \mathcal{D}_2)(r) = \bigwedge_{s \in \text{Tupl}(S)} (\mathcal{D}_2(s) \rightarrow \mathcal{D}_1(rs))$$

for all  $r \in \text{Tupl}(R)$ , where  $\mathcal{D}_1$  is on  $RS$  and  $\mathcal{D}_2$  is on  $S$ .

- ▶ This definition solves the second issue of Codd-style division, but it is domain-dependent. If there is an attribute in  $R$ , that has an infinite domain, the result of division can be infinite.
- ▶ To overcome this issue, some approaches use this definition with additional assumption

$$(\pi_R(\mathcal{D}_1))(r) > 0,$$

which introduces the second issue again.

## Division – Our Approach

- ▶ In our setting,  $(\forall x)\varphi \equiv \neg(\exists x)\neg\varphi$  does not hold in general. We need to consider the division as a fundamental operation.
- ▶ We have proposed domain-independent division, that suffices to establish relational completeness and does not suffer from the second issue of Codd-style division.

### Definition

Let  $\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3$  be RDTs on  $R, S$  and  $RS$ . Division  $\mathcal{D}_1 \div^{\mathcal{D}_3} \mathcal{D}_2$  of  $\mathcal{D}_1$  by  $\mathcal{D}_2$  which ranges over  $\mathcal{D}_3$  is an RDT on  $R$  defined by

$$(\mathcal{D}_1 \div^{\mathcal{D}_3} \mathcal{D}_2)(r) = \bigwedge_{s \in \text{Tupl}(S)} (\mathcal{D}_3(r) \otimes (\mathcal{D}_2(s) \rightarrow \mathcal{D}_1(rs)))$$

for each  $r \in \text{Tupl}(R)$ .

## Date's Small Divide

- ▶ In order to overcome the second issue of Codd-style division, Date proposed division operation called Small Divide. Later, it has been further extended to eliminate the first issue as well.
- ▶ For relations  $\mathcal{D}_1$  on  $R$  (dividend),  $\mathcal{D}_2$  on  $S$  (divisor) and  $\mathcal{D}_3$  on  $RS$  (mediator), the Small Divide can be characterized by

$$\mathcal{D}_1 \div_{\text{sdo}}^{\mathcal{D}_3} \mathcal{D}_2 = \{r \in \mathcal{D}_1 \mid \text{for all } s \in \mathcal{D}_2, \text{ we have } rs \in \mathcal{D}_3\}.$$

- ▶ Its generalization in our model is straightforward:

$$(\mathcal{D}_1 \div_{\text{gsdo}}^{\mathcal{D}_3} \mathcal{D}_2)(r) = \mathcal{D}_1(r) \otimes \bigwedge_{s \in \text{Tupl}(S)} (\mathcal{D}_2(s) \rightarrow \mathcal{D}_3(rs)).$$

- ▶ If the residuated lattice is prelinear or divisible, division in our style is equivalent to generalized Small Divide.

## Todd's Division, Date's Great Divide

- ▶ To overcome the first issue of Codd-style division, in the past Todd proposed the following: For  $\mathcal{D}_1$  on  $RS$  and  $\mathcal{D}_2$  on  $ST$

$$\mathcal{D}_1 \div_{\text{Todd}} \mathcal{D}_2 = \{rt \in \pi_R(\mathcal{D}_1) \bowtie \pi_T(\mathcal{D}_2) \mid \forall s: \text{if } st \in \mathcal{D}_2 \text{ then } rs \in \mathcal{D}_1\}$$

- ▶ This definition suffers from the second issue. Date proposed Great Divide in similar fashion as Small Divide to overcome this issue. For  $\mathcal{D}_1$  on  $R$ ,  $\mathcal{D}_2$  on  $S$ ,  $\mathcal{D}_3$  on  $RS$  and  $\mathcal{D}_4$  on  $ST$ :

$$\mathcal{D}_1 \div_{\text{gdo}}^{\mathcal{D}_3, \mathcal{D}_4} \mathcal{D}_2 = \{rt \in \mathcal{D}_1 \bowtie \mathcal{D}_2 \mid \forall s: \text{if } st \in \mathcal{D}_4 \text{ then } rs \in \mathcal{D}_3\}$$

## Darwen's Divide

- ▶ Darwen's division is based on Great Divide, but do not pose any requirements on relation schemes.
- ▶ For relations  $\mathcal{D}_1$  on  $R_1$  (*dividend*),  $\mathcal{D}_2$  on  $R_2$  (*divisor*),  $\mathcal{D}_3$  on  $R_3$  (*first mediator*), and  $\mathcal{D}_4$  on  $R_4$  (*second mediator*), we put

$$\mathcal{D}_1 \div_{\text{ddo}}^{\mathcal{D}_3, \mathcal{D}_4} \mathcal{D}_2 = (\mathcal{D}_1 \bowtie \mathcal{D}_2) \bar{\bowtie} ((\mathcal{D}_1 \bowtie \mathcal{D}_4) \bar{\bowtie} \mathcal{D}_3).$$

- ▶ Written in set notation:

$$\mathcal{D}_1 \div_{\text{ddo}}^{\mathcal{D}_3, \mathcal{D}_4} \mathcal{D}_2 = \{r_1 r_2 \in \mathcal{D}_1 \bowtie \mathcal{D}_2 \mid$$

for all  $r_4 \in \mathcal{D}_4$ : if  $r_1 r_2 \bar{\bowtie} r_4$ , then there is  $r_3 \in \mathcal{D}_3$ :  $r_1 r_4 \bar{\bowtie} r_3\}$ ,

- ▶ The relation scheme of result of Darwen's Divide is  $R_1 \cup R_2$  since  $R_1$  and  $R_2$  are arbitrary and might have some attributes in common.

## Darwen's Divide – graded variant

- ▶ We may introduce a graded variant  $\div_{\text{gddo}}$  of the Darwen's Divide as follows

$$\begin{aligned}
 (\mathcal{D}_1 \div_{\text{gddo}}^{\mathcal{D}_3, \mathcal{D}_4} \mathcal{D}_2)(r_1 r_2) &= \\
 &= \mathcal{D}_1(r_1) \otimes \mathcal{D}_2(r_2) \otimes \bigwedge_{\substack{r_4 \in \text{Tupl}(R_4) \\ r_1 r_2 \checkmark r_4}} \left( \mathcal{D}_4(r_4) \rightarrow \bigvee_{\substack{r_3 \in \text{Tupl}(R_3) \\ r_1 r_4 \checkmark r_3}} \mathcal{D}_3(r_3) \right) \\
 &= (\mathcal{D}_1 \bowtie \mathcal{D}_2)(r_1 r_2) \otimes \bigwedge_{\substack{r_4 \in \text{Tupl}(R_4) \\ r_1 r_2 \checkmark r_4}} \left( \mathcal{D}_4(r_4) \rightarrow \bigvee_{\substack{r_3 \in \text{Tupl}(R_3) \\ r_1 r_4 \checkmark r_3}} \mathcal{D}_3(r_3) \right)
 \end{aligned}$$

- ▶ Semantics of the operation depends on the relations schemes of input relations.

## Darwen's Divide – graded variant, contd.

- ▶ The condition of joinability is not necessary and can be avoided:

$$\begin{aligned}
 & (\mathcal{D}_1 \dot{\div}_{\text{gddo}}^{\mathcal{D}_3, \mathcal{D}_4} \mathcal{D}_2)(r_1 r_2) = \\
 & = (\mathcal{D}_1 \bowtie \mathcal{D}_2)(r_1 r_2) \otimes \bigwedge_{\substack{r_4 \in \text{Tupl}(R_4) \\ r_1 r_2 \check{\bowtie} r_4}} \left( \mathcal{D}_4(r_4) \rightarrow \bigvee_{\substack{r_3 \in \text{Tupl}(R_3) \\ r_1 r_4 \check{\bowtie} r_3}} \mathcal{D}_3(r_3) \right) \\
 & = (\mathcal{D}_1 \bowtie \mathcal{D}_2)(r_1 r_2) \otimes \bigwedge_{r'_4 \in \text{Tupl}(R_{4 \setminus 12})} \left( \mathcal{D}_4(r_{12}^{\check{\bowtie}4} r'_4) \rightarrow \bigvee_{r'_3 \in \text{Tupl}(R_{3 \setminus 14})} \mathcal{D}_3(r_{14}^{\check{\bowtie}3} r'_3) \right) \\
 & = (\mathcal{D}_1 \bowtie \mathcal{D}_2)(r_1 r_2) \otimes \bigwedge_{r'_4 \in \text{Tupl}(R_{4 \setminus 12})} \left( \mathcal{D}_4(r_{12}^{\check{\bowtie}4} r'_4) \rightarrow \pi_{R_{3 \cap 14}}(\mathcal{D}_3)(r_{14}^{\check{\bowtie}3}) \right),
 \end{aligned}$$

- ▶ where  $r_{12}^{\check{\bowtie}4} = (r_1 r_2)(R_{4 \cap 12})$  and  $r_{14}^{\check{\bowtie}3} = (r_1 r_4)(R_{3 \cap 14})$ .

## Relationship among the operators – easy part

- ▶ If  $\mathbf{L}$  is prelinear or divisible, then division in our sense is equivalent to Date's Small Divide. We have

$$\mathcal{D}_3 \dot{\div}^{\mathcal{D}_1} \mathcal{D}_2 = \mathcal{D}_1 \dot{\div}_{\text{gsdo}}^{\mathcal{D}_3} \mathcal{D}_2$$

- ▶ As in the classical case, Date's Small Divide is expressible by Date's Great Divide in the following way:

$$\mathcal{D}_1 \dot{\div}_{\text{gsdo}}^{\mathcal{D}_3} \mathcal{D}_2 = \mathcal{D}_1 \dot{\div}_{\text{ggdo}}^{\mathcal{D}_3, \mathcal{D}_2} 1_{\emptyset}.$$

- ▶ Darwen's Divide is directly applicable to relations that conform to requirements imposed by Date's Great Divide with the same semantics. We have

$$\mathcal{D}_1 \dot{\div}_{\text{ggdo}}^{\mathcal{D}_3, \mathcal{D}_4} \mathcal{D}_2 = \mathcal{D}_1 \dot{\div}_{\text{gddo}}^{\mathcal{D}_3, \mathcal{D}_4} \mathcal{D}_2.$$



## Relationship among the operators – trickier part

- ▶ In order to show further relationships among the division operations we have introduced new relationally complete query language – Pseudo Tuple Calculus
- ▶ PTC provides easy way to express any relational operation utilizing its set notation.
- ▶ Such PTC expression is semantically equivalent to the original RA operation.
- ▶ Using relational completeness we get the equivalent formulation of this operation utilizing just the fundamental RA operations.

## Relationship among the operators – trickier part

- ▶ Example: Consider the division operation in our sense, i. e., for RDTs  $\mathcal{D}_1$ ,  $\mathcal{D}_2$ , and  $\mathcal{D}_3$  on  $RS$ ,  $S$ , and  $R$ , we put

$$(\mathcal{D}_1 \div^{\mathcal{D}_3} \mathcal{D}_2)(r) = \bigwedge_{s \in \text{Tupl}(S)} (\mathcal{D}_3(r) \otimes (\mathcal{D}_2(s) \rightarrow \mathcal{D}_1(rs))),$$

for each  $r \in \text{Tupl}(R)$ .

- ▶ Consider relation symbols  $\mathbb{D}_1, \mathbb{D}_2$  and  $\mathbb{D}_3$  on  $RS$ ,  $S$ , and  $R$ , respectively. Then the PTC-expression

$$\mathcal{T}(r) = \bigwedge_s (\mathbb{D}_3(r) \otimes (\mathbb{D}_2(s) \rightarrow \mathbb{D}_1(rs))),$$

is semantically equivalent to the division operation. More precisely, for a database instance  $\mathcal{D}$  such that  $\mathbb{D}_1^{\mathcal{D}} = \mathcal{D}_1$ ,  $\mathbb{D}_2^{\mathcal{D}} = \mathcal{D}_2$ , and  $\mathbb{D}_3^{\mathcal{D}} = \mathcal{D}_3$  we have

$$(\mathcal{D}_1 \div^{\mathcal{D}_3} \mathcal{D}_2)(r) = \mathcal{T}^{\mathcal{D}}(r)$$

for all  $r \in \text{Tupl}(R)$ .

## Relationship among the operators – trickier part

Let  $\mathcal{D}_1$ ,  $\mathcal{D}_2$ , and  $\mathcal{D}_3$  be RDTs on  $RS$ ,  $S$ , and  $R$ , respectively, and let  $\div_{\text{gsdo}}$  be Date's Small Divide. For the division operation in our sense we have

$$(\mathcal{D}_1 \div^{\mathcal{D}_3} \mathcal{D}_2)(r) = (\mathcal{E}_R^{\mathcal{D}} \div_{\text{gsdo}}^{E^{\mathcal{D}}} \mathcal{E}_S^{\mathcal{D}})(r),$$

for all  $r \in \text{Tupl}(R)$  where

$$E^{\mathcal{D}} = \mathcal{D}_3 \bowtie ((\mathcal{D}_2 \bowtie \mathcal{E}_R^{\mathcal{D}})^{\mathcal{E}_{RS}^{\mathcal{D}}} \mathcal{D}_1)$$

and the extended active domains  $\mathcal{E}_R^{\mathcal{D}}$ ,  $\mathcal{E}_S^{\mathcal{D}}$ , and  $\mathcal{E}_{RS}^{\mathcal{D}}$  contain tuples built only from the values from relations  $\mathcal{D}_1$ ,  $\mathcal{D}_2$ , and  $\mathcal{D}_3$ .

## Relationship among the operators – trickier part

Let  $\mathcal{D}_1$ ,  $\mathcal{D}_2$ , and  $\mathcal{D}_3$  be RDTs on  $R$ ,  $S$ , and  $RS$ , respectively, and let  $\div$  be the division operation in our sense. For Date's Small Divide we have

$$(\mathcal{D}_1 \div_{\text{gsdo}}^{\mathcal{D}_3} \mathcal{D}_2)(r) = (\mathcal{D}_1 \bowtie (E^{\mathcal{D}} \div_{\mathcal{R}}^{\mathcal{E}^{\mathcal{D}}} \mathcal{E}_S^{\mathcal{D}}))(r),$$

for all  $r \in \text{Tupl}(R)$  where

$$E^{\mathcal{D}} = ((\mathcal{D}_2 \bowtie \mathcal{E}_R^{\mathcal{D}})^{\mathcal{E}_{RS}^{\mathcal{D}}} \mathcal{D}_3)$$

and the extended active domains  $\mathcal{E}_R^{\mathcal{D}}$ ,  $\mathcal{E}_S^{\mathcal{D}}$ , and  $\mathcal{E}_{RS}^{\mathcal{D}}$  contain tuples built only from the values from relations  $\mathcal{D}_1$ ,  $\mathcal{D}_2$ , and  $\mathcal{D}_3$ .

## Relationship among the operators – trickier part

Let  $\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3$ , and  $\mathcal{D}_4$  be RDTs on  $R_1, R_2, R_3$ , and  $R_4$ , respectively, and let  $\div$  be the division operation defined by (??). For Darwen's Divide we have

$$(\mathcal{D}_1 \div_{\text{gddo}}^{\mathcal{D}_3, \mathcal{D}_4} \mathcal{D}_2)(r) = ((\mathcal{D}_1 \bowtie \mathcal{D}_2) \bowtie (E^{\mathcal{D}} \div_{R'_1}^{\mathcal{E}^{\mathcal{D}}} \mathcal{E}_{R'_2}^{\mathcal{D}}))(r),$$

for all  $r \in \text{Tupl}(R_1 \cup R_2)$  where

$$E^{\mathcal{D}} = ((\mathcal{D}_4 \bowtie \mathcal{E}_{R'_3}^{\mathcal{D}})^{\mathcal{E}_{R'_4}^{\mathcal{D}}} (\pi_{R'_3}(\mathcal{D}_3) \bowtie \mathcal{E}_{R'_4}^{\mathcal{D}})),$$

$$R'_1 = (R_4 \cap (R_1 \cup R_2)) \cup (R_1 \cap R_3),$$

$$R'_2 = R_4 \setminus (R_1 \cup R_2),$$

$$R'_3 = R_3 \cap (R_1 \cup R_4),$$

$$R'_4 = R_4 \cup (R_1 \cap R_3)$$

and the extended active domains contain tuples built only from the values from relations  $\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3$ , and  $\mathcal{D}_4$ .

## Note on Graded Subsethood

- ▶ Graded subsethood can be seen as a relational operation on two RDTs  $\mathcal{D}_1, \mathcal{D}_2$  on the same relation scheme  $R$ .
- ▶ The result is an RDT on empty relation scheme containing an empty tuple with score that is equal to the degree of  $\mathcal{D}_1$  being a subset of  $\mathcal{D}_2$ . It is given by

$$(\mathcal{D}_1 \subseteq \mathcal{D}_2)(\emptyset) = \bigwedge_{r \in \text{Tupl}(R)} (\mathcal{D}_1(r) \rightarrow \mathcal{D}_2(r))$$

- ▶ Graded subsethood can be seen as a special case of division in our style, generalized Small Divide, Great Divide or Darwen's Divide.
- ▶ The generalized division operations can be expressed using the graded subsethood and the notion of image relations.

# Conclusions

- ▶ We have presented an overview of various relational division operations and their generalized counterparts.
- ▶ Furthermore, we have investigated their relationship with the help of newly introduced relationally complete query language – Pseudo Tuple Calculus.
- ▶ We have answered an open question regarding the relationship of Date's Great Divide and Darwen's Divide in the classical model.
- ▶ Future research will focus on a model with graded subsethood as a fundamental operation alongside with the imaging operation.