Title: Links to Data Bases and relational algebra

Material: Database System Implementation, Section 6.1.1 to 6.1.5

Database System Concepts, Section 3.2

Chapter 1 of Dechter's manuscript, pages 13 to 17.

Foundations of Constraint Satisfaction

CSCE421/821, Fall 2001

www.cse.unl.edu/~choueiry/F01-421-821/

Berthe Y. Choueiry (Shu-we-ri)

Ferguson Hall, Room 104

choueiry@cse.unl.edu, Tel: (402)472-5444

Warning

There are strong (?) historical and conceptual connections between:

Constraint Databases	$\leftrightarrow$	Constraint Logic Programming
Query processing in Relational Databases	$\leftrightarrow$	Constraint propagation Solving CSPs

- Constraint databases and constraint logic programming share the representation language (FOL)
- Constraint Satisfaction and relational databases share computation mechanisms

Relations

Binary relation: Given two sets D_a and D_b, any set of 2-tuples $\langle x, y \rangle$ with $x \in D_a$ and $y \in D_b$ defines a <u>relation</u> R_{ab}

$$R_{a,b} = \{\,(x, y)\} \subseteq D_a \times D_b$$

Function: *Special binary relation*

For any element x in D_a there is <u>at most one</u> tuple $\langle x, ? \rangle$ in R_{ab}

D_a is called the <u>domain</u> of the the function

D_b is called the <u>range</u> of the function

k-ary relations: Given k sets D_1, D_2, ..., D_k, any set of k-tuples $\langle x_1, x_2, \ldots, x_k \rangle$ with $x_1 \in D_1$, $x_2 \in D_2$, ..., $x_k \in D_k$ defines a <u>relation</u>: $R_{1,2,\ldots,k} = \{\,(x_1, x_2, \ldots, x_k)\} \subseteq D_1 \times D_2 \times \ldots \times D_k$

Representation of relations

Binary arrays:

→ 2-dim binary array (i.e., binary matrix):

$C_{V_k,V_l} = R_{V_k,V_l} = \{\,(x,y)\} \subseteq D_{V_k} \times D_{V_l}$

$$\begin{pmatrix} 1 & 0 & 1 & \ldots & 1 \\ 0 & \ldots & & & 1 \\ 1 & \ldots & & & 1 \end{pmatrix}$$

→ k-dimensional binary arrays

Tables:

C_1		
A	**B**	**V**
2	1	1
4	1	1
6	2	1
1	3	2
2	1	2
4	1	2
6	2	2
2	3	5
4	1	6
6	2	6

C_2	
D	**V**
1	1
1	2
3	3
6	3
4	4
6	4
5	6
6	6

C_3			
E	**F**	**G**	**V**
1	2	1	1
1	2	1	2
2	2	1	2
2	2	3	2
3	2	1	2
1	3	3	6
3	2	1	6

DB terminology	CSP terminology
Table=Relation	Constraint
Relation arity	Constraint arity
Attribute	Variable
Value of an attribute	Value of variable
Domain of an attribute	Domain of a variable
Tuple	Tuple acceptable by a constraint.

Operations on relations:

DB: Intersection, Union, Difference, Selection, Projection, Join (Cartesian product), equi-join, etc.

CSP: + composition

Note:

Intersection, union, difference: relations must have same scope

Selection, projection: defined on one relation (unary operators)

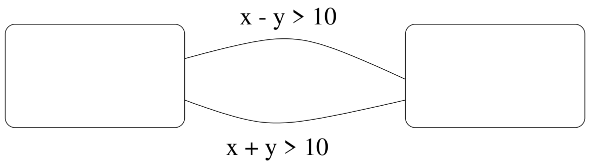
Join: any number of relations with any scope

Intersection

Input: two or more relations of the same scope

Output: a new more restrictive relation with the same scope made of tuples that are in all the input relations (simultaneously)

Matrix operation: logical `AND`



R

x_1	x_2	x_3
a	b	c
b	b	c
c	b	c
c	b	s

R'

x_1	x_2	x_3
b	b	c
c	b	c
c	n	n

R''

x_2	x_3	x_4
a	a	1
b	c	2
b	c	3

$R \cap R'$? $R \cap R''$?

Union

Input: two or more relations of the same scope

Output: a new less <u>restrictive</u> relation with the same scope made of tuples that are in <u>any</u> of the input relations

Matrix operation: logical `OR`

R

x_1	x_2	x_3
a	b	c
b	b	c
c	b	c
c	b	s

R'

x_1	x_2	x_3
b	b	c
c	b	c
c	n	n

R''

x_2	x_3	x_4
a	a	1
b	c	2
b	c	3

$R \cup R'$? $R \cup R''$?

Difference

Input: two relations R and R' of the same scope

Output: a new more restrictive relation than R made of tuples that are in R but not in R'

Matrix operation: Boolean difference

R

x_1	x_2	x_3
a	b	c
b	b	c
c	b	c
c	b	s

R'

x_1	x_2	x_3
b	b	c
c	b	c
c	n	n

R''

x_2	x_3	x_4
a	a	1
b	c	2
b	c	3

$R - R'$? $R - R''$?

Selection

Input: A relation R and some test/predicate on attributes of R

Output: A relation R', same scope than R but with only a subset of the tuples in R (those that satisfy the predicate)

Relation operation: row selection

x_1	x_2	x_3
a	b	c
b	b	c
c	b	c
c	b	s

Select such than $x_1 > x_2$, $\sigma_{x_1 > x_2}(R)$?

Projection

Input: A relation R and a subset s of the scope (attributes)

Output: A relation R' of scope s with the tuples rewritten such than positions not in s are removed

Relation operation: column elimination

x_1	x_2	x_3
a	b	c
b	b	c
c	b	c
c	b	s

Project R on x_1, x_2, $\pi_{x_1,x_2}(R)$?

Join (natural join)

Input: Two relations R and R'

Output: A relation R'', whose scope is union of scopes of R and R' and tuples satisfy both R and R'. $\rightarrow$ R and R' have no attribute common: Cartesian product $\rightarrow$ R and R' have an attribute in common, compute all possibilities

Operation: Compute all solutions to a CSP.

R

x_1	x_2	x_3
a	b	c
b	b	c
c	b	c
c	b	s

R''

x_2	x_3	x_4
a	a	1
b	c	2
b	c	3

Join R and R', $R \bowtie R'$

Composition of relations

[Montanari'74]

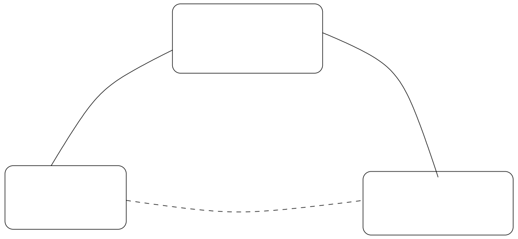
Input: two binary relations R_{ab} and R_{bc} with 1 variable in common

Output: a new induced relation R_{ac} (to be combined by intersection to an pre-existing relation between them, if any).

Matrix operation: binary matrix multiplication

$R_{ac} = R_{ab} \cdot R_{bc} \quad R_{ac,rs} = \bigvee_{t=1}^{D_b} R_{ab,rt} \wedge R_{bc,ts}$

Note: generalization as constraint synthesis (Freuder)



$\longrightarrow$ **Direct vs. induced relations**

$$R_{ab} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \; R_{bc} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix}, \; R_{ac}?$$

Question:

- Given a CSP with three variables V_1, V_2, V_3 and the binary constraints C_{V_1,V_2}, C_{V_2,V_3}, write C_{V_1,V_3}, the composition of these two constraints, in relational algebra.
- Given a CSP with three variables V_1, V_2, V_3 and the binary constraints $C_{V_1,V_2}, C_{V_1,V_3}, C_{V_2,V_3}$, write the new C'_{V_1,V_3} induced between V_1 and V_3, in relational algebra.

DB terminology	CSP terminology
(Natural, inner) Join	Synthesized constraint.
	(Composition/Cartesian product of at least two 2)
Left/right outer join	Synthesized constraint including (some) inconsistent tuples
Projection of a join	Induced constraint
Computing $r_1 \bowtie r_2 \bowtie \ldots \bowtie r_i$	Finding all solutions to the conjunction of constraints $\bigwedge_{k=1}^{i} r_k$
Answering a query	Formulating the constraints and solving a CSP