RELATIONS 3

In this section we will give an application of relations to **databases**. For this we will need the following more general notion of Cartesian products and relations.

Definition (Cartesian product of n sets) : The Cartesian product of n sets; $X_1,\ X_2,\ X_3,\ \ldots,\ X_n$, is denoted by $X_1 \times X_2 \times X_3 \times \ldots \times X_n$ and is defined by:

$$X_1 \times X_2 \times X_3 \times \ldots \times X_n \ = \left\{ \begin{array}{c} \textbf{ordered n-tuples}\ (x_1,\ x_2,\ x_3,\ \ldots,\ x_n) \\ \text{such that} \\ (x_1 \in X_1),\ (x_2 \in X_2),\ (x_3 \in X_3),\ \ldots,\ (x_n \in X_n) \end{array} \right\}.$$

Definition (An n-ary relation) : Given n sets $X_1,\ X_2,\ X_3,\ \ldots,\ X_n$, an n-ary relation on these sets is just a **subset** of the Cartesian product $X_1 \times X_2 \times X_3 \times \ldots \times X_n$.

Example [1] (A ternary relation): Let the sets $X_1,\ X_2,\ X_3,$ be defined as follows:

$$\begin{aligned} X_1 &= \text{All human males.} \\ X_2 &= \text{All human females.} \\ X_3 &= \text{All humans.} \end{aligned}$$

Define a relation $\Re$ on $X_1 \times X_2 \times X_3$ by

$$\boxed{(x_1,\ x_2,\ x_3)\ \in\ \Re} \iff \boxed{\begin{array}{l} \qquad x_1 \ \text{is the father of}\ x_3 \\ \text{and} \\ \qquad x_2 \ \text{is the mother of}\ x_3 \end{array}}\ .$$

Example [2] (A quaternary relation): Let $X_1 = X_2 = X_3 = X_4 = \mathbb{Z}$ and define a relation $\Re$ on $X_1 \times X_2 \times X_3 \times X_4$ by

$$\boxed{(x_1,\ x_2,\ x_3,\ x_4)\ \in\ \Re} \iff \boxed{x_1 x_4 \ =\ x_2 x_3}\ .$$

Example [3] (A n-ary relation): Consider the sets

$$X_1 = X_2 = X_3 = \ldots = X_n = \{\text{All humans}\}$$

and define a relation $\Re$ on $X_1 \times X_2 \times X_3 \times \ldots \times X_n$ by

$$\boxed{(x_1, x_2, x_3, \ldots, x_{n-1}, x_n) \in \Re} \iff \boxed{\begin{array}{c} x_1 \text{ is a parent of } x_2, \\ x_2 \text{ is a parent of } x_3, \\ x_3 \text{ is a parent of } x_4, \\ \vdots \\ x_{n-1} \text{ is a parent of } x_n \end{array}}.$$

Note that such a relation might be considered if one wanted to study family trees.

Relations are used to structure databases in computing as follows: We begin by defining a set of **attributes**

$$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \ldots, \boldsymbol{\alpha}_n\},$$

which the objects in some set $\boldsymbol{\mathcal{S}}$ may possess. For example, if **$\boldsymbol{\mathcal{S}}$ is the set of all humans**, then

attribute $\boldsymbol{\alpha}_1$ might be the persons name,
attribute $\boldsymbol{\alpha}_2$ might be the persons age (measured in full years)
attribute $\boldsymbol{\alpha}_3$ might be the persons current address

$\vdots$

To each attribute $\boldsymbol{\alpha}_j$ we specify a corresponding set, X_j say. For example,

corresponding to attribute $\boldsymbol{\alpha}_1$ we could specify that $X_1 = \{\text{all human names}\}$
corresponding to attribute $\boldsymbol{\alpha}_2$ we could specify that $X_2 = \mathbb{N}$
corresponding to attribute $\boldsymbol{\alpha}_3$ we could specify that $X_3 = \{\text{all addresses}\}$

$\vdots$

Definition (Relational database) : Let $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \ldots, \boldsymbol{\alpha}_n$ be a set of attributes with corresponding sets $X_1, X_2, X_3, \ldots, X_n$. A relational database with attributes $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \ldots, \boldsymbol{\alpha}_n$ is just set of relations, $\{\Re_1, \Re_2, \Re_3 \ldots, \Re_D\}$, where each relation $\Re_k$ is a relation between some (possibly all) of the sets $X_1, X_2, X_3, \ldots, X_n$. These relations, $\{\Re_1, \Re_2, \Re_3 \ldots, \Re_D\}$ are called record files.

Example [4] (A relational database): Here is a relational database consisting of two record files, File 1 and File 2 given below, that a bank might have. The attributes in this example are:

attribute $\boldsymbol{\alpha}_1 =$ Name,
attribute $\boldsymbol{\alpha}_2 =$ Account Number
attribute $\boldsymbol{\alpha}_3 =$ Deposit Amount
attribute $\boldsymbol{\alpha}_4 =$ Withdrawl Amount
attribute $\boldsymbol{\alpha}_5 =$ Date
attribute $\boldsymbol{\alpha}_6 =$ Balance
attribute $\boldsymbol{\alpha}_7 =$ Address

The corresponding sets $X_1, X_2, X_3, \ldots, X_7$ are

$X_1 = \{\text{all names}\}$
$X_2 = \mathbb{N}$
$X_3 = \{\text{positive real numbers}\}$
$X_4 = \{\text{positive real numbers}\}$
$X_5 = \{\text{triples of two digit non-negative integers}\}$
$X_6 = \mathbb{R}$
$X_7 = \{\text{all addresses}\}$

Name	Account Number	Deposit Amount	Withdrawl Amount	Date	Balance
Adams, T.	55003321	104.23		16-09-08	437.19
Adams, T.	55003321	24.23		30-09-08	461.42
Adams, T.	55003321		100.00	08-10-08	361.42
Adams, T.	55003321	140.00		12-10-08	501.42
Egan, M.	96144305		500.00	03-09-08	-472.67
Egan, M.	96144305	100.00		24-09-08	-372.67
Egan, M.	96144305		200.00	07-10-08	-572.67
Murphy, J.	68305265		150.00	15-09-08	-843.43
Murphy, J.	68305265	200.00		06-10-08	-643.43

File 1.

Name	Account Number	Address
Adams, T.	55003321	1 The Hill, Dublin 9
Egan, M.	96144305	44 Montenotte, Cork
Murphy, J.	68305265	6 The Claddagh, Galway

File 2.

File 1 $\subseteq X_1 \times X_2 \times X_3 \times X_4 \times X_6 \times X_6$
File 2 $\subseteq X_1 \times X_2 \times X_7$

New Databases from Old : If we start with a relational database (such as that given above which consists of the two files, File 1 and File 2) we can construct new record files by performing various operations on the existing files. The most important of these operations are project, select and join which we now define:

Definition (project): For a relational database with attributes $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \ldots, \boldsymbol{\alpha}_n$ we form new record files, denoted by

$$\text{Project}\,[\,\text{File Name},\ (\boldsymbol{\alpha}_{j_1}, \boldsymbol{\alpha}_{j_2}, \boldsymbol{\alpha}_{j_3}, \ldots, \boldsymbol{\alpha}_{j_m}\,)\,]$$

by deleting all **columns** from " File Name " except those specified by given attributes $(\boldsymbol{\alpha}_{j_1}, \boldsymbol{\alpha}_{j_2}, \boldsymbol{\alpha}_{j_3}, \ldots, \boldsymbol{\alpha}_{j_m}\,)$

Example [5] (project): The file, File 3, given by

$$\text{File 3} = \text{Project}\,[\,\text{File 1},\ (\,\text{Name},\ \text{Account Number},\ \text{Date},\ \text{Balance}\,)\,]$$

is the following:

Name	Account Number	Date	Balance
Adams, T.	55003321	16-09-08	437.19
Adams, T.	55003321	30-09-08	461.42
Adams, T.	55003321	08-10-08	361.42
Adams, T.	55003321	12-10-08	501.42
Egan, M.	96144305	03-09-08	-472.67
Egan, M.	96144305	24-09-08	-372.67
Egan, M.	96144305	07-10-08	-572.67
Murphy, J.	68305265	15-09-08	-843.43
Murphy, J.	68305265	06-10-08	-643.43

File 3.

That is, we just delete from File 1 the columns headed "Deposit Amount" and "Withdrawl Amount".

Definition (select): For a relational database with attributes $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \ldots, \boldsymbol{\alpha}_n$ we form new record files, denoted by

$$\text{Select}\,[\,\text{File Name},\ \text{Criterion}\,]$$

by deleting all **rows** from " File Name " except those specified by the given Criterion.

Example [6] (select): The file, File 4, given by

$$\text{File 4} = \text{Select}\,[\,\text{File 1},\ (\,\text{Date after 30-09-08},\ \text{Balance} < 0\ )\,]$$

is the following:

Name	Account Number	Deposit Amount	Withdrawl Amount	Date	Balance
Egan, M.	96144305		200.00	07-10-08	-572.67
Murphy, J.	68305265	200.00		06-10-08	-643.43

File 4.

That is, we just delete from File 1 all rows except the third last row and the last row.

Terminology (Type of a record file)): For a relational database with attributes $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \ldots, \boldsymbol{\alpha}_n$ we say that record file, $\boldsymbol{\mathcal{F}}$, is of type $(\boldsymbol{\alpha}_{j_1}, \boldsymbol{\alpha}_{j_2}, \ldots, \boldsymbol{\alpha}_{j_m})$ if and only if

$$\boldsymbol{\mathcal{F}} \subseteq X_{j_1} \times X_{j_2} \times \ldots \times X_{j_m}$$

where X_i is the set associated with the attribute $\boldsymbol{\alpha}_i$. For example;

$$\text{File 3 is of type}\ (\,\text{Name},\ \text{Account Number},\ \text{Date},\ \text{Balance}\,)$$

Definition (join): Let $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_\ell\,;\ \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_m\,;\ \gamma_1, \gamma_2, \ldots, \gamma_n$ be attributes with corresponding sets: $X_1, X_2, \ldots X_\ell\,;\ Y_1, Y_2, \ldots, Y_m\,;\ Z_1, Z_2, \ldots, Z_n$.

If $\boldsymbol{\mathcal{F}}$ is a file of type $(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_\ell, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_m)$ and
if $\boldsymbol{\mathcal{G}}$ is a file of type $(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_\ell, \gamma_1, \gamma_2, \ldots, \gamma_n)$

we define the file join $(\boldsymbol{\mathcal{F}}, \boldsymbol{\mathcal{G}})$, called the join of $\boldsymbol{\mathcal{F}}$ and $\boldsymbol{\mathcal{G}}$, to be a file of type;

$$(\,\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_\ell\,;\ \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_m\,;\ \gamma_1, \gamma_2, \ldots, \gamma_n\,)$$

such that

$$(x_1, x_2, \ldots, x_\ell, y_1, y_2, \ldots, y_m, z_1, z_2, \ldots, z_n) \in \text{join}(\boldsymbol{\mathcal{F}}, \boldsymbol{\mathcal{G}})$$
$$\Updownarrow$$
$$(x_1, x_2, \ldots, x_\ell, y_1, y_2, \ldots, y_m) \in \boldsymbol{\mathcal{F}} \text{ and } (x_1, x_2, \ldots, x_\ell, z_1, z_2, \ldots, z_n) \in \boldsymbol{\mathcal{G}}$$

Note : $(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_\ell)$ are called **the common attributes** of files $\boldsymbol{\mathcal{F}}$ and $\boldsymbol{\mathcal{G}}$. Though we have listed here as the first ℓ attributes of both $\boldsymbol{\mathcal{F}}$ and $\boldsymbol{\mathcal{G}}$, generally this need not be the case. In fact, we can even perform the join when $\ell = 0$, that is when $\boldsymbol{\mathcal{F}}$ and $\boldsymbol{\mathcal{G}}$ have no attributes in common at all.

Example [7] (join): File 5 = join (File 2, File 3) is given by:

Name	Account Number	Address	Date	Balance
Adams, T.	55003321	1 The Hill, Dublin 9	16-09-08	437.19
Adams, T.	55003321	1 The Hill, Dublin 9	30-09-08	461.42
Adams, T.	55003321	1 The Hill, Dublin 9	08-10-08	361.42
Adams, T.	55003321	1 The Hill, Dublin 9	12-10-08	501.42
Egan, M.	96144305	44 Montenotte, Cork	03-09-08	-472.67
Egan, M.	96144305	44 Montenotte, Cork	24-09-08	-372.67
Egan, M.	96144305	44 Montenotte, Cork	07-10-08	-572.67
Murphy, J.	68305265	6 The Claddagh, Galway	15-09-08	-843.43
Murphy, J.	68305265	6 The Claddagh, Galway	06-10-08	-643.43

File 5.

Here the attributes are:

$$\boldsymbol{\alpha}_1 = \text{Name}, \qquad \boldsymbol{\alpha}_2 = \text{Account Number},$$

$$\boldsymbol{\beta}_1 = \text{Address},$$

$$\boldsymbol{\gamma}_1 = \text{Date}, \qquad \boldsymbol{\gamma}_2 = \text{Balance}.$$