CPSC 304 – Administrative notes October 9 & October 10, 2024

- Project:
  - Milestone 1 is returned
    - There will be a comment from me/your project TA. Any comment from a TA is from your project TA
  - October 15: Milestone 2 due
  - Project repositories are ready (see piazza)
  - Oracle accounts are ready (see Piazza)
  - Tutorial materials have been released to help you
    - Oracle/SQL Plus
    - PHP
    - JavaScript
- October 22: Midterm @ 6PM
  - Those who registered conflicts have been notified of the time for that exam. if you registered conflicts and can't make that time or did not register a conflict, please see the policy in the syllabus.

CPSC 304
Introduction to Database Systems

Formal Relational Languages

Textbook Reference
Database Management Systems: 4 - 4.2
(skip the calculii)

Learning Goals

- Identify the basic operators in Relational Algebra (RA).
- Use RA to create queries that include combining RA operators.
- Given an RA query and table schemas and instances, compute the result of the query.

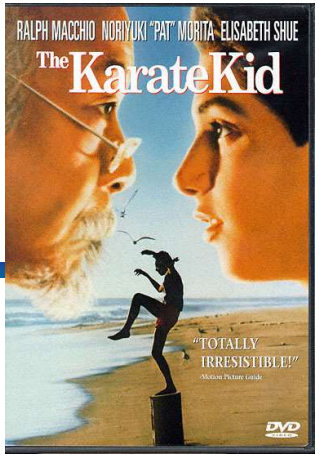
Databases: the continuing saga

When last we left databases…

- We learned that they're excellent things
- We learned how to conceptually model them using ER diagrams
- We learned how to logically model them using relational schemas
- We knew how to normalize our database relations

We're almost ready to use SQL to query it, but first…

Balance, Daniel-san, is key



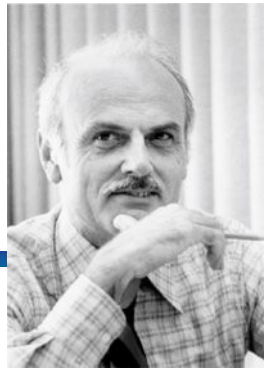
The mathematical foundations:

- Relational Algebra
  - Clear way of describing core concepts
  - *partially procedural*: describe what you want and how you want it, but the order of operations matters

Relational Query Languages

- Allow data manipulation and retrieval from a DB
- Relational model supports simple, powerful QLs:
  - Strong formal foundation based on logic
  - Allows for much optimization via *query optimizer*
- Query Languages **!=** Programming Languages
  - QLs not intended for complex calculations
  - QLs provide *easy access* to large datasets
  - Users *do not* need to know how to navigate through complicated data structures

Relational Algebra (RA) All in one place



- Basic operations:
  - *Selection* (σ): Selects a subset of rows from relation.
  - *Projection* (π): Deletes unwanted columns from relation.
  - *Cross-product* (x): Allows us to combine two relations.
  - *Set-difference* (-): Tuples in relation 1, but not in relation 2.
  - *Union* ($\cup$): Tuples in relation 1 and in relation 2.
  - *Rename* (ρ): Assigns a (another) name to a relation
- Additional, inessential but useful operations:
  - *Intersection* ($\cap$), *join* ($\bowtie$), *division* (/), *assignment*($\leftarrow$)
- All operators take one or two relations as inputs and give a new relation as a result
- For the purposes of relational algebra, relations are sets
- Operations can be ***composed***. (Algebra is "closed")

Example Movies Database

Movie(MovieID, Title, Year)

StarsIn(MovieID, StarID, Character)

MovieStar(StarID, Name, Gender)

Example Instances

Movie:

MovieID	Title	Year
1	Star Wars	1977
2	Casablanca	1942
3	The Wizard of Oz	1939
4	Indiana Jones and the Raiders of the Lost Ark	1981

StarsIn:

MovieID	StarID	Character
1	1	Han Solo
4	1	Indiana Jones
2	2	Ilsa Lund
3	3	Dorothy Gale

MovieStar:

StarID	Name	Gender
1	Harrison Ford	Male
2	Ingrid Bergman	Female
3	Judy Garland	Female

Selection (σ (sigma))

- Notation: $\sigma_p(r)$
- p is called the **selection predicate**
- Defined as:

$$\sigma_p(r) = \{t \mid t \in r \textbf{ and } p(t)\}$$

Set of tuples of r satisfying p

Where p is a formula in propositional calculus consisting of:

connectives : $\wedge$ (**and**), $\vee$ (**or**), $\neg$ (**not**)

and

predicates:

<attribute> *op* <attribute> or

<attribute> *op* <constant>

where *op* is one of: =, ≠, >, ≥, <, ≤

Selection Example

Movie:

MovieID	Title	Year
1	Star Wars	1977
2	Casablanca	1942
3	The Wizard of Oz	1939
4	Indiana Jones and the Raiders of the Lost Ark	1981

$\sigma_{year > 1945}(Movie)$

MovieID	Title	Year
1	Star Wars	1977
4	Indiana Jones and the Raiders of the Lost Ark	1981

Selection Example

Find all male stars from the MovieStar table.

StarID	Name	Gender
1	Harrison Ford	Male

$$\sigma_{\text{Gender = 'Male'}}\text{MovieStar}$$

Projection (π (pi))

- Notation:

  $\pi_{A1, A2, ..., Ak}(r)$

  where *A1, …,Ak* are attributes (the projection list) and *r* is a relation.
- The result: a relation of the *k* attributes A1, A2, …, AK obtained from r by erasing the columns that are not listed
- Duplicate rows removed from result (relations are sets)

Projection Examples

Movie:

MovieID	Title	Year
1	Star Wars	1977
2	Casablanca	1942
3	The Wizard of Oz	1939
4	Indiana Jones and the Raiders of the Lost Ark	1981

$\pi_{\text{Title, Year}}$ (Movie)

Title	Year
Star Wars	1977
Casablanca	1942
The Wizard of Oz	1939
Indiana Jones and the Raiders of the Lost Ark	1981

π_{Year}(Movie)

Year
1977
1939
1942
1981

What is $\pi_{\text{Title,Year}}(\sigma_{\text{year} > 1945}(\text{Movie}))$?

Title	Year
Star Wars	1977
Indiana Jones and the Raiders of the Lost Ark	1981

Projection Example #2

Find the IDs of actors who have starred in movies

$$\pi_{\text{StarID}}(\text{StarsIn})$$

StarID
1
2
3

Clicker Projection Example

Suppose relation R(A,B,C) has the tuples:

A	B	C
1	2	3
4	2	3
4	5	6
2	5	3
1	2	6

Compute the projection $\pi_{C,B}(R)$, and identify one of its tuples from the list below.

A. (2,3)
B. (4,2,3)
C. (6,4)
D. (6,5)
E. None of the above

Clicker Projection Example

Suppose relation R(A,B,C) has the tuples:

A	B	C
1	2	3
4	2	3
4	5	6
2	5	3
1	2	6

Compute the projection $\pi_{C,B}(R)$, and identify one of its tuples from the list below.

A. (2,3) Wrong order

B. (4,2,3) Not projected

C. (6,4) Wrong attributes

D. (6,5) right

E. None of the above

C	B
3	2
6	5
3	5
6	2

Selection and Projection Example

Find the ids of movies made prior to 1950

Movie:

MovieID	Title	Year
1	Star Wars	1977
2	Casablanca	1942
3	The Wizard of Oz	1939
4	Indiana Jones and the Raiders of the Lost Ark	1981

MovieID
2
3

Selection and Projection Example

Find the ids of movies made prior to 1950

Movie:

MovieID	Title	Year
1	Star Wars	1977
2	Casablanca	1942
3	The Wizard of Oz	1939
4	Indiana Jones and the Raiders of the Lost Ark	1981

$$\pi_{\text{MovieID}} (\sigma_{\text{year} < 1950} (\text{Movie}))$$

MovieID
2
3

Selection and Projection Example

Would this work?

A) Yes

B) No

MovieID	Title	Year
1	Star Wars	1977
2	Casablanca	1942
3	The Wizard of Oz	1939
4	Indiana Jones and the Raiders of the Lost Ark	1981

$$\sigma_{\text{year} < 1950}(\pi_{\text{MovieID}}(\text{Movie}))$$

MovieID
2
3

Selection and Projection Example

Would this work?

A) Yes

B) No

MovieID	Title	Year
1	Star Wars	1977
2	Casablanca	1942
3	The Wizard of Oz	1939
4	Indiana Jones and the Raiders of the Lost Ark	1981

$$\sigma_{\text{year} < 1950}(\pi_{\text{MovieID}}(\text{Movie}))$$

MovieID
2
3

Union, Intersection, Set-Difference

- Notation: $\boldsymbol{r \cup s}$ $\boldsymbol{r \cap s}$ $\boldsymbol{r - s}$
- Defined as:

  $r \cup s = \{t \mid t \in r \text{ or } t \in s\}$
  $r \cap s = \{ t \mid t \in r \textbf{ and } t \in s \}$
  $r - s = \{t \mid t \in r \textbf{ and } t \notin s\}$
- For these operations to be well-defined:
  1. *r, s* must have the *same arity* (same number of attributes)
  2. The attribute domains must be *compatible* (e.g., 2nd column of *r* has same domain of values as the 2nd column of *s*)
- What is the schema of the result?

Union, Intersection, and Set Difference Examples

MovieStar

StarID	Name	Gender
1	Harrison Ford	Male
2	Ingrid Bergman	Female
3	Judy Garland	Female

Singer

StarID	SName	Gender
3	Judy Garland	Female
4	Sam Smith	Non-binary

MovieStar ∪ Singer

StarID	Name	Gender
1	Harrison Ford	Male
2	Ingrid Bergman	Female
3	Judy Garland	Female
4	Sam Smith	Non-binary

MovieStar ∩ Singer

StarID	Name	Gender
3	Judy Garland	Female

MovieStar – Singer

StarID	Name	Gender
1	Harrison Ford	Male
2	Ingrid Bergman	Female

Set Operator Example

MovieStar

StarID	Name	Gender
1	Harrison Ford	Male
2	Ingrid Bergman	Female
3	Judy Garland	Female

Singer

StarID	Name	Gender
3	Judy Garland	Female
4	Sam Smith	Non-binary

Find the names of stars that are Singers but not MovieStars

Name
Sam Smith

Set Operator Example

MovieStar

StarID	Name	Gender
1	Harrison Ford	Male
2	Ingrid Bergman	Female
3	Judy Garland	Female

Singer

StarID	Name	Gender
3	Judy Garland	Female
4	Sam Smith	Non-binary

Find the names of stars that are Singers but not MovieStars

$\pi_{\text{Name}}(\text{Singer} - \text{MovieStar})$

Name
Sam Smith

Cartesian (or Cross)-Product

- Notation: $r \times s$
- Defined as:

$$r \times s = \{ t \; q \mid t \in r \textbf{ and } q \in s \}$$

- It is possible for r and s to have attributes with the same name, which creates a naming conflict.
  - In this case, the attributes are referred to solely by position.

Cartesian Product Example

MovieStar

StarID	Name	Gender
1	Harrison Ford	Male
2	Ingrid Bergman	Female
3	Judy Garland	Female

StarsIn

MovieID	StarID	Character
1	1	Han Solo
4	1	Indiana Jones
2	2	Ilsa Lund
3	3	Dorothy Gale

MovieStar x StarsIn

1	Name	Gender	MovieID	5	Character
1	Harrison Ford	Male	1	1	Han Solo
2	Ingrid Bergman	Female	1	1	Han Solo
3	Judy Garland	Female	1	1	Han Solo
1	Harrison Ford	Male	4	1	Indiana Jones
2	Ingrid Bergman	Female	4	1	Indiana Jones
3	Judy Garland	Female	4	1	Indiana Jones
…	…	…	…	…	…

Rename (ρ (rho))

- Allows us to name results of relational-algebra expressions.
- Notation

  $$\rho\ (X, E)$$

  returns the expression *E* under the name *X*

- We can rename part of an expression, e.g.,
  $\rho((StarID \rightarrow ID), \pi_{StarID,Name}(MovieStar))$

- We can also refer to positions of attributes, e.g.,
  $\rho((1 \rightarrow ID), \pi_{StarID,Name}(MovieStar))$
  Is the same as above

Rename (ρ (rho))

- We can rename the resulting relation and the attributes in that relation

$$\rho(GenderlessStars(ID,Nom), \pi_{StarID,Name}(MovieStar))$$

MovieStar

StarID	Name	Gender
1	Harrison Ford	Male
2	Ingrid Bergman	Female
3	Judy Garland	Female

$\pi_{StarID,Name}$(MovieStar)

StarID	Name
1	Harrison Ford
2	Ingrid Bergman
3	Judy Garland

GenderlessStars

ID	Nom
1	Harrison Ford
2	Ingrid Bergman
3	Judy Garland

ρ Example

MovieStar

StarID	Name	Gender
1	Harrison Ford	Male
2	Ingrid Bergman	Female
3	Judy Garland	Female

StarsIn

MovieID	StarID	Character
1	1	Han Solo
4	1	Indiana Jones
2	2	Ilsa Lund
3	3	Dorothy Gale

ρ((1→StarID1, 5→StarID2), MovieStar x StarsIn)

StarID1	Name	Gender	MovieID	StarID2	Character
1	Harrison Ford	Male	1	1	Han Solo
2	Ingrid Bergman	Female	1	1	Han Solo
3	Judy Garland	Female	1	1	Han Solo
1	Harrison Ford	Male	4	1	Indiana Jones
2	Ingrid Bergman	Female	4	1	Indiana Jones
3	Judy Garland	Female	4	1	Indiana Jones
…	…	…	…	…	…

Additional Operations

- They can be defined in terms of the primitive operations
- They are added for convenience
- They are:
  - Join (Condition, Equi-, Natural) ($\bowtie$)
  - Division (/)
  - Assignment ($\leftarrow$)

Joins ($\bowtie$)

- *Condition Join*:

$$R \bowtie_c S = \sigma_c(R \times S)$$

- *Result schema* same as cross-product.
- Fewer tuples than cross-product
  - might be able to compute more efficiently
- Sometimes called a *theta-join*.
  - The reference to an attribute of a relation R can be by position (R.i) or by name (R.name)

Condition Join Example

MovieStar

StarID	Name	Gender
1	Harrison Ford	Male
2	Ingrid Bergman	Female
3	Judy Garland	Female

StarsIn

MovieID	StarID	Character
1	1	Han Solo
4	1	Indiana Jones
2	2	Ilsa Lund
3	3	Dorothy Gale

MovieStar $\bowtie_{\text{MovieStar.StarID} < \text{StarsIn.StarID}}$ StarsIn

1	Name	Gender	MovieID	5	Character
1	Harrison Ford	Male	2	2	Ilsa Lund
1	Harrison Ford	Male	3	3	Dorothy Gale
2	Ingrid Bergman	Female	3	3	Dorothy Gale

MovieStar $\bowtie_{\text{MovieStar.StarID} < \text{StarsIn.StarID}}$ StarsIn

MovieStar x StarsIn (first get the cross product)

1	Name	Gender	MovieID	5	Character
1	Harrison Ford	Male	1	1	Han Solo
2	Ingrid Bergman	Female	1	1	Han Solo
3	Judy Garland	Female	1	1	Han Solo
1	Harrison Ford	Male	4	1	Indiana Jones
2	Ingrid Bergman	Female	4	1	Indiana Jones
3	Judy Garland	Female	4	1	Indiana Jones
1	Harrison Ford	Male	2	2	Ilsa Lund
2	Ingrid Bergman	Female	2	2	Ilsa Lund
3	Judy Garland	Female	2	2	Ilsa Lund
1	Harrison Ford	Male	3	3	Dorothy Gale
2	Ingrid Bergman	Female	3	3	Dorothy Gale
3	Judy Garland	Female	3	3	Dorothy Gale

MovieStar ⋈ MovieStar.StarID < StarsIn.StarID StarsIn

Now remove rows based on the condition stated above.

1	Name	Gender	MovieID	5	Character
~~1~~	~~Harrison Ford~~	~~Male~~	~~1~~	~~1~~	~~Han Solo~~
~~2~~	~~Ingrid Bergman~~	~~Female~~	~~1~~	~~1~~	~~Han Solo~~
~~3~~	~~Judy Garland~~	~~Female~~	~~1~~	~~1~~	~~Han Solo~~
~~1~~	~~Harrison Ford~~	~~Male~~	~~4~~	~~1~~	~~Indiana Jones~~
~~2~~	~~Ingrid Bergman~~	~~Female~~	~~4~~	~~1~~	~~Indiana Jones~~
~~3~~	~~Judy Garland~~	~~Female~~	~~4~~	~~1~~	~~Indiana Jones~~
1	Harrison Ford	Male	2	2	Ilsa Lund
~~2~~	~~Ingrid Bergman~~	~~Female~~	~~2~~	~~2~~	~~Ilsa Lund~~
~~3~~	~~Judy Garland~~	~~Female~~	~~2~~	~~2~~	~~Ilsa Lund~~
1	Harrison Ford	Male	3	3	Dorothy Gale
2	Ingrid Bergman	Female	3	3	Dorothy Gale
~~3~~	~~Judy Garland~~	~~Female~~	~~3~~	~~3~~	~~Dorothy Gale~~

Condition Join Clicker Example

- Compute $R \bowtie_{R.A < S.C \text{ and } R.B < S.D} S$ where:

R(A,B):

A	B
1	2
3	4
5	6

S(B,C,D):

B	C	D
2	4	6
4	6	8
4	7	9

Assume the schema of the result is (A, R.B, S.B, C, D).
Which tuple is in the result?

A. (1,2,2,6,8)
B. (1,2,4,4,6)
C. (5,6,2,4,6)
D. All are valid
E. None are valid

Condition Join Clicker Example

- Compute $R \bowtie_{R.A < S.C \text{ and } R.B < S.D} S$ where:

R(A,B):

A	B
1	2
3	4
5	6

S(B,C,D):

B	C	D
2	4	6
4	6	8
4	7	9

Assume the schema of the result is (A, R.B, S.B, C, D).

Which tuple is in the result?

A. (1,2,2,6,8) (2,6,8) would have to be in S

B. (1,2,4,4,6) (4,4,6) would have to be in S

C. (5,6,2,4,6) Violates R.A < SC & R.B < S.D (5 > 4, and 6 = 6)

D. All are valid

E. None are valid Correct

Equi-Join & Natural Join

- *Equi-Join*: A special case of condition join $R \bowtie_c S = \sigma_c(R \times S)$, where c contains only ***equalities***. Only the first of the repeated columns is retained.
- *Natural Join*: Equijoin on ***all*** common attributes
  - *Result schema:* similar to cross-product, but has only one copy of each common attribute
  - No need to show the condition
  - If the two attributes have no common attributes, this would be the same as cross product.
  - This is what we saw in BCNF & 3NF

Equi-Join Example

MovieStar

StarID	Name	Gender
1	Harrison Ford	Male
2	Ingrid Bergman	Female
3	Judy Garland	Female

StarsIn

MovieID	StarID	Character
1	1	Han Solo
4	1	Indiana Jones
2	2	Ilsa Lund
3	3	Dorothy Gale

MovieStar $\bowtie_{\text{MovieStar.StarID} = \text{StarsIn.StarID}}$ StarsIn

StarID	Name	Gender	MovieID	Character
1	Harrison Ford	Male	1	Han Solo
1	Harrison Ford	Male	4	Indiana Jones
2	Ingrid Bergman	Female	2	Ilsa Lund
3	Judy Garland	Female	3	Dorothy Gale

Natural Join Examples

MovieStar

StarID	Name	Gender
1	Harrison Ford	Male
2	Ingrid Bergman	Female
3	Judy Garland	Female

StarsIn

MovieID	StarID	Character
1	1	Han Solo
4	1	Indiana Jones
2	2	Ilsa Lund
3	3	Dorothy Gale

MovieStar ⋈ StarsIn

StarID	Name	Gender	MovieID	Character
1	Harrison Ford	Male	1	Han Solo
1	Harrison Ford	Male	4	Indiana Jones
3	Judy Garland	Female	3	Dorothy Gale
2	Ingrid Bergman	Female	2	Ilsa Lund

Join Example

Movie(MovieID, Title, Year)
StarsIn(MovieID, StarID, Character)
MovieStar(StarID, Name, Gender)

- Find the names of all Movie Stars who were in any Movie

Name
Harrison Ford
Ingrid Bergman
Judy Garland

Join Example

Movie(MovieID, Title, Year)
StarsIn(MovieID, StarID, Character)
MovieStar(StarID, Name, Gender)

- Find the names of all Movie Stars who were in any Movie

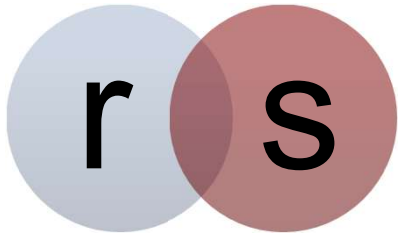
$$\pi_{name}(\text{MovieStar} \bowtie \text{StarsIn})$$

Name
Harrison Ford
Ingrid Bergman
Judy Garland

One more thing you may find helpful: Assignment Operation

- Notation: $t \leftarrow E$
  assigns the result of expression E to a temporary relation t.
- Used to break complex queries to small steps.
- Assignment is always made to a temporary relation variable.
- Example: Write $r \cap s$ in terms of $\cup$ and/or –

$$temp1 \leftarrow r - s$$
$$result \leftarrow r - temp1$$



Okay, let’s do some exercises!

Find names of actors who have been in 'Indiana Jones'

(note: examples use 'Indiana Jones and Raiders of the Lost Ark', but no one wants to write all that out)

$$(\sigma_{\text{Title = 'Indiana Jones'}} \text{Movie})$$

MovieID	Title	Year
4	Indiana Jones and the Raiders of the Lost Ark	1981

$$((\sigma_{\text{Title = 'Indiana Jones'}} \text{Movie}) \bowtie \text{StarsIn})$$

MovieID	Title	Year	StarID	Character
4	Indiana Jones and the Raiders of the Lost Ark	1981	1	Indiana Jones

$$(\pi_{\text{Name}}((\sigma_{\text{Title = 'Indiana Jones'}} \text{Movie}) \bowtie \text{StarsIn} \bowtie \text{MovieStar}))$$

Name
Harrison Ford

Find the name of actors who have been in 'Indiana Jones' and 'Star Wars'

$$Indy \leftarrow \pi_{starID}((\sigma_{Title = 'Indiana\ Jones'} Movie) \bowtie StarsIn)$$

$$StarWars \leftarrow \pi_{starID}((\sigma_{Title = 'Star\ Wars'} Movie) \bowtie StarsIn)$$

$$CoolPeople \leftarrow Indy \cap StarWars$$

$$\pi_{name}(CoolPeople \bowtie MovieStar)$$

(note: examples use 'Indiana Jones and Raiders of the Lost Ark', but no one wants to write all that out)

Find the name of actors who have been in 'Indiana Jones' and 'Star Wars'

Would this work?

$$Indy \leftarrow \pi_{name}((\sigma_{Title = 'Indiana\ Jones'} Movie) \bowtie StarsIn \bowtie MovieStar)$$

$$StarWars \leftarrow \pi_{name}((\sigma_{Title = 'Star\ Wars'} Movie) \bowtie StarsIn \bowtie MovieStar)$$

$$CoolPeople \leftarrow Indy \cap StarWars$$

(note: examples use 'Indiana Jones and Raiders of the Lost Ark', but no one wants to write all that out)

Find the name of actors who have been in 'Indiana Jones' and 'Star Wars'

MovieStar(StarID, Name, Gender)

StarID	Name	Gender
1	Harrison Ford	Male
2	Ingrid Bergman	Female
3	Judy Garland	Female
4	Harrison Ford	Male

StarsIn(MovieID, StarID, Character)

MovieID	StarID	Character
1	4	Han Solo
4	1	Indiana Jones
2	2	Ilsa Lund
3	3	Dorothy Gale

Movie(MovieID, Title, Year)

MovieID	Title	Year
1	Star Wars	1977
2	Casablanca	1942
3	The Wizard of Oz	1939
4	Indiana Jones and the Raiders of the Lost Ark	1981

Exercise

Find the names of actors who have been in a movie with the same title as the actor's name

Clicker Exercise

Find the names of actors who have been in a movie with the same title as the actor's name

Which of the following does *not* do that correctly:

A. $\pi_{\text{Name}}((\text{StarsIn} \bowtie \text{Movie}) \bowtie_{\text{Name = Title } \wedge \text{ StarID = MovieStar.StarID}} \text{MovieStar})$

B. $\pi_{\text{Name}}(\text{MovieStar} \bowtie_{\text{Name = Title } \wedge \text{ MovieStar.StarID = StarID}} (\text{StarsIn} \bowtie \text{Movie}))$

C. $\pi_{\text{Name}}((\text{StarsIn} \bowtie (\pi_{\text{StarID,Name}} \text{MovieStar})) \bowtie_{\text{MovieID = Movie.MovieID } \wedge \text{ Title = Name}} \text{Movie})$

D. All are correct

E. None are correct

All are correct (D)

Clicker Exercise Explained Option A

$\pi_{\text{Name}}((\text{StarsIn} \bowtie \text{Movie}) \bowtie_{\text{Name} = \text{Title} \wedge \text{StarID} = \text{MovieStar.StarID}} \text{MovieStar})$

StarsIn ⋈ Movie

Movie ID	StarID	Character	Title	Year
1	1	Han Solo	Star Wars	1977
2	2	Ilsa Lund	Casablanca	1942
3	3	Dorothy Gale	The Wizard of Oz	1939
4	1	Indiana Jones	Indiana Jones and the Raiders of the Lost Ark	1981

Clicker Exercise Explained Option A

$\pi_{\text{Title}}((\text{Movie} \bowtie \text{StarsIn}) \bowtie_{\text{Name} = \text{Title} \wedge \text{StarID} = \text{MovieStar.StarID}} \text{MovieStar})$

Why do we need *StarID* = *MovieStar.StarID*?

What does *Name* = *Title* do?

Let's process this like a cross product and then remove tuples that don't fit our condition.

(StarsIn⋈ Movie) x MovieStar

Movie ID	2	Character	Title	Year	6	Name	Gender
1	1	Han Solo	Star Wars	1977	1	Harrison Ford	Male
2	2	Ilsa Lund	Casablanca	1942	1	Harrison Ford	Male
3	3	Dorothy Gale	The Wizard of Oz	1939	1	Harrison Ford	Male
4	1	Indiana Jones	Indiana Jones and the Raiders of the Lost Ark	1981	1	Harrison Ford	Male
1	1	Han Solo	Star Wars	1977	2	Ingrid Bergman	Female
2	2	Ilsa Lund	Casablanca	1942	2	Ingrid Bergman	Female
3	3	Dorothy Gale	The Wizard of Oz	1939	2	Ingrid Bergman	Female
4	1	Indiana Jones	Indiana Jones and the Raiders of the Lost Ark	1981	2	Ingrid Bergman	Female
1	1	Han Solo	Star Wars	1977	3	Judy Garland	Female
2	2	Ilsa Lund	Casablanca	1942	3	Judy Garland	Female
3	3	Dorothy Gale	The Wizard of Oz	1939	3	Judy Garland	Female
4	1	Indiana Jones	Indiana Jones and the Raiders of the Lost Ark	1981	3	Judy Garland	Female

Remove tuples where starIDs (columns 4 and 6) do not match

Movie ID	2	Character	Title	Year	6	Name	Gender
1	1	Han Solo	Star Wars	1977	1	Harrison Ford	Male
2	2	Ilsa Lund	Casablanca	1942	1	Harrison Ford	Male
3	3	Dorothy Gale	The Wizard of Oz	1939	1	Harrison Ford	Male
4	1	Indiana Jones	Indiana Jones and the Raiders of the Lost Ark	1981	1	Harrison Ford	Male
1	1	Han Solo	Star Wars	1977	2	Ingrid Bergman	Female
2	2	Ilsa Lund	Casablanca	1942	2	Ingrid Bergman	Female
3	3	Dorothy Gale	The Wizard of Oz	1939	2	Ingrid Bergman	Female
4	1	Indiana Jones	Indiana Jones and the Raiders of the Lost Ark	1981	2	Ingrid Bergman	Female
1	1	Han Solo	Star Wars	1977	3	Judy Garland	Female
2	2	Ilsa Lund	Casablanca	1942	3	Judy Garland	Female
3	3	Dorothy Gale	The Wizard of Oz	1939	3	Judy Garland	Female
4	1	Indiana Jones	Indiana Jones and the Raiders of the Lost Ark	1981	3	Judy Garland	Female

Remove tuples where starIDs (columns 4 and 6) do not match

Movie ID	2	Character	Title	Year	6	Name	Gender

Remove tuples where starIDs (columns 2 and 6) do not match

Movie ID	2	Character	Title	Year	6	Name	Gender
1	1	Han Solo	Star Wars	1977	1	Harrison Ford	Male
4	1	Indiana Jones	Indiana Jones and the Raiders of the Lost Ark	1981	1	Harrison Ford	Male
2	2	Ilsa Lund	Casablanca	1942	2	Ingrid Bergman	Female
3	3	Dorothy Gale	The Wizard of Oz	1939	3	Judy Garland	Female

Only keep tuples where Name = Title

Movie ID	2	Character	Title	Year	6	Name	Gender

π_{Name}

Name

Note: Based on this relational instance, there were no results but this does not mean the relational algebra expression is incorrect.

Clicker Exercise Explained Option B

π_{Name}(MovieStar $\bowtie_{\text{Name = title ^ MovieStar.StarID = StarID}}$ (StarsIn $\bowtie$ Movie))

The same join as option A but the order of the joins is different.

Clicker Exercise Explained Option C

$\pi_{Name}((StarsIn \bowtie (\pi_{StarID,Name} MovieStar)) \bowtie_{MovieID = Movie.MovieID \wedge title = name} Movie)$

$\pi_{StarID,Name} MovieStar$

StarID	Name
1	Harrison Ford
2	Ingrid Bergman
3	Judy Garland

$StarsIn \bowtie (\pi_{StarID,Name} MovieStar)$

MovieID	StarID	Character	Name
1	1	Han Solo	Harrison Ford
4	1	Indiana Jones	Harrison Ford
2	2	Ilsa Lund	Ingrid Bergman
3	3	Dorothy Gale	Judy Garland

Clicker Exercise Explained Option C

$\pi_{Name}((StarsIn \bowtie (\pi_{StarID,Name} MovieStar)) \bowtie_{MovieID = Movie.MovieID \wedge title = name} Movie)$

$$StarsIn \bowtie (\pi_{StarID,Name} MovieStar)$$

MovieID	StarID	Character	Name
1	1	Han Solo	Harrison Ford
4	1	Indiana Jones	Harrison Ford
2	2	Ilsa Lund	Ingrid Bergman
3	3	Dorothy Gale	Judy Garland

Same as before. Let's do the cross product and eliminate tuples that do not meet our listed condition.

Movie ID	StarID	Character	Name	5	Title	Year
1	1	Han Solo	Harrison Ford	1	Star Wars	1977
4	1	Indiana Jones	Harrison Ford	1	Star Wars	1977
2	2	Ilsa Lund	Ingrid Bergman	1	Star Wars	1977
3	3	Dorothy Gale	Judy Garland	1	Star Wars	1977
1	1	Han Solo	Harrison Ford	2	Casablanca	1942
4	1	Indiana Jones	Harrison Ford	2	Casablanca	1942
2	2	Ilsa Lund	Ingrid Bergman	2	Casablanca	1942
3	3	Dorothy Gale	Judy Garland	2	Casablanca	1942
1	1	Han Solo	Harrison Ford	3	The Wizard of Oz	1939
4	1	Indiana Jones	Harrison Ford	3	The Wizard of Oz	1939
2	2	Ilsa Lund	Ingrid Bergman	3	The Wizard of Oz	1939
3	3	Dorothy Gale	Judy Garland	3	The Wizard of Oz	1939
1	1	Han Solo	Harrison Ford	4	Indiana Jones and the Raiders of the Lost Ark	1981
4	1	Indiana Jones	Harrison Ford	4	Indiana Jones and the Raiders of the Lost Ark	1981
2	2	Ilsa Lund	Ingrid Bergman	4	Indiana Jones and the Raiders of the Lost Ark	1981
3	3	Dorothy Gale	Judy Garland	4	Indiana Jones and the Raiders of the Lost Ark	1981

1	StarID	Character	Name		Title	Year
1	1	Han Solo	Harrison Ford		Star Wars	1977
4	1	Indiana Jones	Harrison Ford		Star Wars	1977
2	2	Ilsa Lund	Ingrid Bergman		Star Wars	1977
3	3	Dorothy Gale	Judy Garland		Star Wars	1977
1	1	Han Solo	Harrison Ford		Casablanca	1942
4	1	Indiana Jones	Harrison Ford		Casablanca	1942
2	2	Ilsa Lund	Ingrid Bergman		Casablanca	1942
3	3	Dorothy Gale	Judy Garland		Casablanca	1942
1	1	Han Solo	Harrison Ford		The Wizard of Oz	1939
4	1	Indiana Jones	Harrison Ford		The Wizard of Oz	1939
2	2	Ilsa Lund	Ingrid Bergman		The Wizard of Oz	1939
3	3	Dorothy Gale	Judy Garland		The Wizard of Oz	1939
1	1	Han Solo	Harrison Ford		Indiana Jones and the Raiders of the Lost Ark	1981
4	1	Indiana Jones	Harrison Ford		Indiana Jones and the Raiders of the Lost Ark	1981
2	2	Ilsa Lund	Ingrid Bergman		Indiana Jones and the Raiders of the Lost Ark	1981
3	3	Dorothy Gale	Judy Garland		Indiana Jones and the Raiders of the Lost Ark	1981

Only keep tuples where title = name

MovieID	StarID	Character	Name	Title	Year
1	1	Han Solo	Harrison Ford	Star Wars	1977
2	2	Ilsa Lund	Ingrid Bergman	Casablanca	1942
3	3	Dorothy Gale	Judy Garland	The Wizard of Oz	1939
4	1	Indiana Jones	Harrison Ford	Indiana Jones and the Raiders of the Lost Ark	1981

MovieID	StarID	Character	Name	Year

Projection only keeps the Name attribute

Name