Algebra and Join Minimization

John Clara, Danyang Zhang

Winter 2016.

How to Optimize Queries

- Perform different mappings to reduce rows
- Answer variables cannot change
- Constants cannot change
- Everything else is fair game!

Example 1

What are all the books by the person who wrote ”Twilight”?

Example 1

What are all the books by the person who wrote ”Twilight”?

```
SELECT b1.title
FROM Book b1, Book b2, Book b3
WHERE b1.author = b2.author AND
b3.author = b2.author AND
b3.title = ”Twilight”;
```

Example 1

What are all the books by the person who wrote ”Twilight”?

```
SELECT b1.title
FROM Book b1, Book b2, Book b3
WHERE b1.author = b2.author AND
b3.author = b2.author AND
b3.title = ”Twilight”;
```

Book	title	author
	d	a
	-	a
	”Twilight”	a

answer	title
	d

Can we map first row to any rows?

Example 1

What are all the books by the person who wrote ”Twilight”?

Book	title	author
	d	a
	-	a
	”Twilight”	a

answer	title
	d

Map second row to some row?

Example 1

What are all the books by the person who wrote ”Twilight”?

Book	title	author
	d	a
	”Twilight”	a

answer	title
	d

Map second row to some row?

Example 1

What are all the books by the person who wrote ”Twilight”?

Book	title	author
	d	a
	”Twilight”	a

answer	title
	d

```
SELECT b1.title
FROM Book b1, Book b2
WHERE b1.author = b2.author AND
b2.title = ”Twilight”;
```

Example 2

```
SELECT t1.A, t2.B, t4.C
FROM R t1, R t2, R t3, R t4, R t5
WHERE t3.A=t4.A AND
t2.B=t3.B AND
t1.C=t2.C AND
t3.C=t5.C AND
t3.A=t5.A;
```

Example 2

```
SELECT t1.A, t2.B, t4.C
FROM R t1, R t2, R t3, R t4, R t5
WHERE t3.A=t4.A AND
t2.B=t3.B AND
t1.C=t2.C AND
t3.C=t5.C AND
t3.A=t5.A;
```

R	A	B	C
	a	-	c1
	-	b	c1
	a1	b	c2
	a1	-	c
	a1	-	c2

answer	A	B	C
	a	c	c

Example 2

R	A	B	C
	a	-	c1
	-	b	c1
	a1	b	c2
	a1	-	c
	a1	-	c2

answer	A	B	C
	a	b	c

Can we reduce any rows?

Example 2

R	A	B	C
	a	-	c1
	-	b	c1
	a1	b	c2
	a1	-	c

answer	A	B	C
	a	b	c

Example 2

R	A	B	C
	a	-	c1
	-	b	c1
	a1	b	-
	a1	-	c

answer	A	B	C
	a	b	c

How to Chase

- For any dependency $X \rightarrow A$, if any tuple agrees on X make them agree on A.

Example 2

Dependencies: $F = \{AC \rightarrow B, B \rightarrow C, C \rightarrow A\}$

R	A	B	C
	a	-	c1
	-	b	c1
	a1	b	-
	a1	-	c

answer	A	B	C
	a	b	c

Example 2

Dependencies: $F = \{AC \to B, B \to C, C \to A\}$

Use $B \to C$

R	A	B	C
	a	-	c1
	-	b	c1
	a1	b	-
	a1	-	c

answer	A	B	C
	a	b	c

Example 2

Dependencies: $F = \{AC \to B, B \to C, C \to A\}$

Use $C \to A$

R	A	B	C
	a	-	c1
	-	b	c1
	a1	b	c1
	a1	-	c

answer	A	B	C
	a	b	c

Example 2

Dependencies: $F = \{AC \rightarrow B, B \rightarrow C, C \rightarrow A\}$

Eliminate rows

R	A	B	C
	a	-	c1
	a	b	c1
	a	b	c1
	a	-	c

answer	A	B	C
	a	b	c

Example 2

Dependencies: $F = \{AC \rightarrow B, B \rightarrow C, C \rightarrow A\}$

Can we use any Dependencies?

R	A	B	C
	a	b	c1
	a	-	c

answer	A	B	C
	a	b	c

Example 2

Dependencies: $F = \{AC \rightarrow B, B \rightarrow C, C \rightarrow A\}$

R	A	B	C
	a	b	-
	a	-	c

answer	A	B	C
	a	b	c

```
SELECT r1.A, r1.B, r2.C
FROM R r1, R r2
WHERE r1.a = r2.a;
```

Reference

1. "*Database Systems Concepts*" by Silberschatz, Korth and Sudarshan, 6th edition, McGraw-Hill.