

Theta Join

- A join that involves a predicate

$$R1 \bowtie_{\theta} R2 = \sigma_{\theta} (R1 \times R2)$$

- Here θ can be any condition

Note that equi-join is a special case of theta join where θ is an equality condition

Examples of Theta-Joins (1)

- Find all hospitals within 5 miles of a school

$\Pi_{\text{name}}(\text{Hospitals} \bowtie_{\text{distance}(\text{location}, \text{location}) < 5} \text{Schools})$

```
SELECT DISTINCT h.name  
FROM Hospitals h, Schools s  
WHERE distance(h.location, s.location) < 5
```

Examples of Theta-Joins (2)

- Find all user clicks made within 5 seconds of a page load

```
SELECT *  
FROM Clicks c, PageLoads p  
WHERE abs(c.click_time - p.load_time) < 5
```

$\text{Clicks} \bowtie_{\text{abs}(\text{click_time} - \text{load_time}) < 5} \text{PageLoads}$

- You might hear “band join” or “range join”

More Joins

- Outer join
 - Include tuples with no matches in the output
 - Use NULL values for missing attributes

- Variants

- Left outer join X
- Right outer join X
- Full outer join X

R1 X R2
— —

Outer Join Example

AnonPatient P

<u>age</u>	<u>zip</u>	disease
54	<u>98125</u>	heart
20	98120	flu
33	98120	lung

AnonJob J

job	<u>age</u>	<u>zip</u>
lawyer	<u>54</u>	<u>98125</u>
cashier	20	98120

P ⋈ V

age	zip	disease	job
54	98125	heart	lawyer
20	98120	flu	cashier
33	98120	lung	null