

COMP 543: Tools & Models for Data Science

Lecture 3 Relational Calculus Handout

FREQUENTS (DRINKER, CAFE)

SERVES (CAFE, COFFEE)

FREQUENTS

DRINKER	CAFE
Chris	A
Chris	B
Chris	C
Risa	A
Risa	B

SERVES

CAFE	COFFEE
A	Drip
A	Cold Brew
A	Espresso
B	Drip
C	Espresso

We want to know:

Who has not gone to a cafe serving Cold Brew?

Earlier, we established that we can answer the question:

Who has gone to a cafe that serves Cold Brew?

with

$$\{f.DRINKER \mid \text{FREQUENTS}(f) \wedge \exists(s, l)(\text{SERVES}(s) \\ \wedge s.CAFE = f.CAFE \\ \wedge s.COFFEE = \text{'Cold Brew'})\}$$

That might lead us to think that we could answer:

Who has not gone to a cafe serving Cold Brew?

With

$$\{f.DRINKER \mid \text{FREQUENTS}(f) \wedge \neg \exists(s, l)(\text{SERVES}(s) \\ \wedge s.CAFE = f.CAFE \\ \wedge s.COFFEE = \text{'Cold Brew'})\}$$

However,

$$\{f.DRINKER \mid \text{FREQUENTS}(f) \wedge \neg \exists(s)(\text{SERVES}(s) \wedge s.CAFE = f.CAFE \wedge s.COFFEE = \text{'Cold Brew'})\}$$

Actually returns: **Who has gone to a cafe that does not serve Cold Brew?**

For convenience, let's say:

$$\text{hasCB} = \text{SERVES}(s) \wedge s.CAFE = f.CAFE \wedge s.COFFEE = \text{'Cold Brew'}$$

Let's walk through this expression.

Consider only the data for Risa

1. Start with the FREQUENTS table
2. Then look at matches in the SERVES table, where s.CAFE = f.CAFE
3. Evaluate the predicate
4. Is it TRUE?
 - If Yes, then include f.DRINKER in the result set
5. When f.CAFE = 'B', Risa gets included in the result set.

f.DRINKER	f.CAFE	s.CAFE	s.COFFEE	hasCB	¬hasCB	Result Set
Risa	B	B	Drip	F	T	{Risa}

6. When f.CAFE = 'A', Risa does NOT get add to the result set

f.DRINKER	f.CAFE	s.CAFE	s.COFFEE	hasCB	¬hasCB	Result Set
Risa	A	A	Drip	T	F	{ }
		A	Cold Brew			
		A	Espresso			

- To get the final result set, we union together all the results from the final column

						{ Risa }
--	--	--	--	--	--	----------

- The final result set is $\{ \text{Risa} \} \cup \{ \} = \{ \text{Risa} \}$
- However, Risa shouldn't be in the result set, because she frequents Cafe B, which serves Cold Brew
- Issue: We want to look at ALL of the coffees served at ALL of the cafes Risa frequents all at one time

Who has not gone to a cafe serving Cold Brew?

To answer this question, we need to introduce a second variable:

$$\{f_1.DRINKER \mid FREQUENTS(f_1) \wedge \neg \exists(f_2, s)(FREQ(f_2) \wedge SERVES(s) \wedge f_2.CAFE = s.CAFE \wedge s.COFFEE = 'Cold Brew' \wedge f_1.DRINKER = f_2.DRINKER)\}$$

Again, for convenience, let's say:

$$hasCB = (FREQ(f_2) \wedge SERVES(s) \wedge f_2.CAFE = s.CAFE \wedge s.COFFEE = 'Cold Brew' \wedge f_1.DRINKER = f_2.DRINKER)\}$$

In this case, by having the second variable, we are able to look at all the data for every place Risa frequents as a whole.

- Here, we have another variable, f_2
- We consider each drinker in turn from the FREQUENTS relation. Basically, we are using this table as our master list of drinkers, and are ignoring the CAFE attribute.

Again, look just at Risa.

DRINKER
Risa

- Now, look at all the combinations of FREQUENTS and SERVES where the CAFE matches and the drinker is $f_1.DRINKER$

$f_1.DRINKER$	$f_2.DRINKER$	$f_2.CAFE$	s.CAFE	s.COFFEE	hasCB	¬hasCB	Result Set
Risa	Risa	A	A	Cold Brew	T	F	{ }
	Risa	A	A	Drip			
	Risa	A	A	Espresso			
	Risa	B	B	Drip			

- If there is any tuple where the Coffee is 'Cold Brew', we exclude the drinker
- Now, in this case, one of the cafes that Risa frequents does serve Cold Brew, so Risa is not added to the result set