Finding a Rewriting

**Theorem** Given views $V_1, \ldots, V_n$ and query $Q$, the problem whether $Q$ has a complete rewriting in terms of $V_1, \ldots, V_n$ is NP complete.
Certain Answers

• Sometimes we can’t answer, but we can get close

\[
\begin{align*}
V1(x,y) & : - E(x,u), E(u,v), E(v,y) \\
V2(x,y) & : - E(x,u), E(u,y), \text{Black}(y)
\end{align*}
\]

\[
Q(x) : - E(x,u), E(u,v), E(v,w), E(w,s)
\]

Can’t really answer Q, but we can find approximations….
Certain Answers

V1(x,y) :- E(x,u), E(u,v), E(v,y)
V2(x,y) :- E(x,u), E(u,y), Black(y)

Q(x) :- E(x,u), E(u,v), E(v,w), E(w,s)

Q(x) :- V2(x,u), V2(u,v)
Q(x) :- V1(x,u), V2(u,v)
Q(x) :- V1(x,u), V1(u,v)

All these return ‘certain’ answers…
Certain Answers

Definition. Given $V_1, \ldots, V_n$, their answers $A_1, \ldots, A_n$ and a query $Q$, a tuple $t$ is a certain tuple for $Q$ iff for every database instance $D$:

- if $A_1 = V_1(D)$ and $\ldots$ and $A_n = V_n(D)$ then $t \in Q(D)$
  
  CWD (Closed World Assumption)

- if $A_1 \subseteq V_1(D)$ and $\ldots$ and $A_n \subseteq V_n(D)$ then $t \in Q(D)$

  OWD (Open World Assumption)
Computing Certain Answers Under OWD

\[
\begin{align*}
V1(x, y) & :\ E(x, u), \ E(u, v), \ E(v, y) \\
V2(x, y) & :\ E(x, u), \ E(u, y), \ Black(y)
\end{align*}
\]

\[
\begin{align*}
Q(x) & :\ E(x, u), \ E(u, v), \ E(v, w), \ E(w, s)
\end{align*}
\]

\[
\begin{align*}
E(x, f(x, y)) & :\ V1(x, y) \\
E(f(x, y), g(x, y)) & :\ V1(x, y) \\
E(g(x, y), y) & :\ V1(x, y) \\
E(x, h(x, y)) & :\ V2(x, y) \\
E(h(x, y), y) & :\ V2(x, y) \\
Black(y) & :\ V2(x, y)
\end{align*}
\]
Computing Certain Answers Under OWD

Next, we have two options

• Run the combined “datalog” program
  – It is actually a Prolog program
  – Notice: data complexity is PTIME

• Transform the datalog program first, so Q returns only values that are not Skolem Terms
Computing Answer Under CWD
Is Different

V1(x) :- R(x,u)
V2(y) :- R(v,y)
Q(x,y) :- R(x,y)

A1 = {a}
A2 = {b}

Certain answers for Q under OWD: none
Certain answers for Q under CWD: (a,b)

Why?