3. Datalog and Limitations of First-Order Logic

1. Datalog Queries over Graphs

We are given two directed graphs G_{black} and G_{white} over the same set V of vertices, represented as binary relations. Write datalog programs that compute

1. the set of pairs $\langle a, b\rangle$ of vertices such that there exists a path from a to b,
2. the set of pairs $\langle a, b\rangle$ of vertices such that there exists a path from a to b, where black and white edges alternate, starting with a white edge.

2. Datalog Queries over a Flight Database

Suppose in a travel agency database there is a table with the schema

$$\texttt{flight}(\texttt{from}, \texttt{to}, \texttt{airline}),$$

where an entry $(c1,\ c2,\ l)$ means that it is possible to fly from city $c1$ to city $c2$ with the airline l.
Consider the following two queries:

1. Return all pairs of cities (x, y) such that is possible to travel from x to y using flights of a *single* airline.
2. Return all pairs of cities (x, y) such that is possible to travel from x to y using flights of *no more than two* airlines.

Which of these queries can you express in datalog?

3. Transitive Closure and First-order Logic

Is it possible to express transitive closure in first-order logic? In other words, given a binary relation $R(\cdot, \cdot)$, is it possible to write a formula $\varphi(x, y)$ such that in every interpretation of R, the formula φ is satisfied exactly by those pairs of domain elements that are in the transitive closure R^+ of R?
Hint: Consider the following statements about a point a and their formulation as logical formulas:

- There exists a point reachable from a that is an R-sink, that is, no R-edge is emanating from that point.
- For each natural number k: All points reachable from a via a path of length k have an emanating R-edge.