Need some kind of pattern expression language

- ## Relational Algebra

Assume a relation R(subject, predicate, object)

$$\sigma_{p=interferes\_with}(R) \bowtie_{o=s} \sigma_{p=regulates}(R) \bowtie_{o=s} \sigma_{p=associated\_with}(R)$$

Need some kind of pattern expression language

- Datalog, 2nd formulation

Assume relations

InterferesWith(drug1, drug2),
Regulates(drug, gene),
AssociatedWith(gene, disease)

Ans(x) :-
InterferesWith(x, y),
Regulates(y, z),
AssociatedWith(z, w)

Need some kind of pattern expression language

- ## SQL, 2nd formulation

Assume relations
> InterferesWith(drug1, drug2),
> Regulates(drug, gene),
> AssociatedWith(gene, disease)

SELECT i.drug1
FROM InterferesWith i, Regulates r, AssociatedWith a
WHERE i.drug2 = r.drug
AND r.gene = a.gene

Aside on SPARQL

- Designed for graph query

- Not algebraically closed
  - Input is a graph, but output is a set of records

- Limited expressiveness
  - Support for path expressions since v1.1
  - But not arbitrary recursion

- If your input is tabular, you have to shred it into a graph before using SPARQL
  - Often a 3x-6x blow up in size!