Logic Programming

Query Evaluation

Michael Genesereth

Computer Science Department

Stanford University
Dataset: \{p(b), p(c), p(d), q(d)\}

Rule:

\[ \text{goal}(X) := p(X) \land \lnot q(X) \]

Instances:

\[ \text{goal}(a) := p(a) \land \lnot q(a) \]
\[ \text{goal}(b) := p(b) \land \lnot q(b) \]
\[ \text{goal}(c) := p(c) \land \lnot q(c) \]
\[ \text{goal}(d) := p(d) \land \lnot q(d) \]

Result: \{\text{goal}(b), \text{goal}(c)\}
Query Semantics

Dataset: \( \{ p(b), p(c), p(d), q(d) \} \)

Rule:

\[
\text{goal}(f(X)) \leftarrow p(X) \land \neg q(X)
\]

Instances:

\[
\begin{align*}
\text{goal}(a) & \leftarrow p(a) \land \neg q(a) \\
\text{goal}(f(a)) & \leftarrow p(f(a)) \land \neg q(f(a)) \\
\text{goal}(f(f(a))) & \leftarrow p(f(f(a))) \land \neg q(f(f(a)))
\end{align*}
\]

... 

Result: \( \{ \text{goal}(b), \text{goal}(c) \} \)
Programme

Evaluation Procedure
  Evaluating ground queries
  Matching
  Evaluating queries with variables

Computational Analysis
  Unindexed datasets
  Indexing
Evaluating Ground Queries
Evaluation of Ground Queries

Given a query rule, call the procedure \texttt{eval} on the body. The result is a boolean. The result is the singleton set of the head if true; else, the empty set.

\textbf{Dataset:} \{p(a,b), p(a,c), p(b,c), p(c,d)\}

\textbf{Query:} goal(c) :- p(c,d) & \neg p(d,c)
\textbf{Body:} p(c,d) & \neg p(d,c)
\textbf{Result:} true

\textbf{Answer:} \{goal(c)\}
(1) If the body of a query rule is an atom, we check whether that atom is contained in our dataset. If so, the body is true.

**Dataset:** \{p(a,b), p(a,c), p(b,c), p(c,d)\}

**Query:** goal(a) :- p(a,b)

**Result:** \{goal(a)\}

**Dataset:** \{p(a,b), p(a,c), p(b,c), p(c,d)\}

**Query:** goal(a) :- p(b,a)

**Result:** \{\}
(2) If the body is a **negation**, we check whether the atom is contained in our dataset. If so, the body is false. If the atom is not contained in our dataset, then the body is true.

**Dataset**: \{p(a,b), p(a,c), p(b,c), p(c,d)}

**Query**: \(\text{goal}(b) :- \sim p(b,c)\)

**Result**: \{\}

**Dataset**: \{p(a,b), p(a,c), p(b,c), p(c,d)}

**Query**: \(\text{goal}(b) :- \sim p(c,b)\)

**Result**: \{\text{goal}(b)\}
(3) If the body is a **conjunction** of literals, we execute this procedure on the first conjunct. If the answer is true, we move on to the next conjunct and so forth until we are done. If the answer to any one of the conjuncts is false, then the value of the body as a whole is false.

**Dataset:** \{p(a,b), p(a,c), p(b,c), p(c,d)\}

**Query:** goal(c) :- p(c,d) & ~p(d,c)

**Result:** \{goal(c)\}

**Dataset:** \{p(a,b), p(a,c), p(b,c), p(c,d)\}

**Query:** goal(c) :- p(c,d) & p(d,c)

**Result:** {}
The value of a query with multiple rules is the union of the values of each of the rules in the query.

**Dataset:** \{p(a,b), p(a,c), p(b,c), p(c,d)\}

**Query:**

- \texttt{goal(a) :- p(a,b)}
- \texttt{goal(b) :- \neg p(b,c)}
- \texttt{goal(c) :- p(c,d) & \neg p(d,c)}

**Result:** \{\texttt{goal(a)}\} \cup \{\} \cup \{\texttt{goal(c)}\}

- \{\texttt{goal(a), goal(c)}\}
Matching
**Matching** is the process of determining whether a *pattern* (an expression with or without variables) matches an instance (an expression without variables), i.e. whether the two expressions can be made identical by appropriate substitutions for the variables in the pattern.
A substitution is a finite set of pairs of variables and terms, called replacements.

\[ \{X \leftarrow a, Y \leftarrow b\} \]

The result of applying a substitution \( \sigma \) to an expression \( \phi \) is the expression \( \phi \sigma \) obtained from \( \phi \) by replacing every occurrence of every variable with a binding in the substitution by the term to which it is bound.

\[
\begin{align*}
  p(X,b)\{X \leftarrow a, Y \leftarrow b\} &= p(a,b) \\
  q(X,Y,X)\{X \leftarrow a, Y \leftarrow b\} &= q(a,b,a)
\end{align*}
\]
A substitution $\sigma$ is a *matcher* for a pattern and an instance if and only if applying the substitution to the pattern results in the given instance.

\[
\begin{align*}
p(X, b)\{X\leftarrow a, Y\leftarrow b\} &= p(a, b) \\
q(X, Y, X)\{X\leftarrow a, Y\leftarrow b\} &= q(a, b, a)
\end{align*}
\]

Here, $\{X\leftarrow a, Y\leftarrow b\}$ is a matcher for $p(X, b)$ and $p(a, b)$. It is also a matcher for $q(X, Y, X)$ and $q(a, b, a)$. 
Matching Procedure

(1) If the pattern is a symbol and the instance is the same symbol, then the procedure succeeds, returning the unmodified substitution as result. If the pattern is a symbol and the instance is a different symbol or a compound expression, then the procedure fails.

(2) If the pattern is a variable with a binding, we compare the binding for the variable with the given instance. If they are identical, the procedure succeeds, returning the unmodified substitution as result; otherwise it fails. If the pattern is a variable without a binding, we include a binding for the variable in the given instance and we return that substitution as a result.

(3) If the pattern is a compound expression and the instance is a compound expression of the same length, we iterate across the pattern and the instance. If the pattern is a compound expression and the instance is a symbol or a compound expression of a different length, we fail.
Example

Compare: \( p(X, Y), p(a, b), {} \)
  Compare: \( p, p, {} \)
  Result: \{\}
  Compare: \( X, a, {} \)
  Result: \{X ← a\}
  Compare: \( Y, b, \{X ← a\} \)
  Result: \{X ← a, Y ← b\}
Result: \{X ← a, Y ← b\}
Example

Compare: \( p(x, x), p(a, a), \{\} \)
  Compare: \( p, p, \{\} \)
  Result: \( \{\} \)
  Compare: \( x, a, \{\} \)
  Result: \( \{x \leftarrow a\} \)
  Compare: \( x, a, \{x \leftarrow a\} \)
    Compare: \( a, a, \{x \leftarrow a\} \)
    Result: \( \{x \leftarrow a\} \)
  Result: \( \{x \leftarrow a\} \)
Result: \( \{x \leftarrow a\} \)
Example

Compare: \( p(x, x), p(a, b), \emptyset \)
  
  Compare: \( p, p, \emptyset \)
  
  Result: \( \emptyset \)

Compare: \( x, a, \emptyset \)
  
  Result: \( \{x \leftarrow a\} \)

Compare: \( x, b, \{x \leftarrow a\} \)
  
  Compare: \( a, b, \{x \leftarrow a\} \)
  
  Result: \( false \)

Result: \( false \)

Result: \( false \)

Result: \( false \)
Evaluation with Variables
Given a query rule, call the procedure `eval` (to be described) on the body and an empty substitution. The result is a list of substitutions that satisfy the body. The value of the rule is obtained by applying the substitutions to the head of the rule.

**Dataset:** \{p(a,b), p(a,c), p(b,c), p(c,d)\}

**Query:** \texttt{goal(Y) :- p(a,Y) \& p(Y,Z)}

Call eval: \(p(a,Y) \& p(Y,Z), \Theta\)

Exit eval: \(\{Y\leftarrow b, Z\leftarrow c\}, \{Y\leftarrow c, Z\leftarrow d\}\)

Value of query: \(\{\text{goal(b), goal(c)}\}\)
(1) If the expression is an **atom**, we try matching the atom to the factoids in our dataset. For each factoid that matches the atom, we add the corresponding substitution to our answer set; and we return the set of all substitutions obtained in this way.

**Dataset**: \{p(a,b), p(a,c), p(b,c), p(c,d)\}

**Call eval**: p(a,Y), {}
**Exit eval**: \{Y←b\}, \{Y←c\}
(2) If the expression is a **negation**, we call eval on the target of the negation and the given substitution. If the result is a non-empty set, then the negation is false and we return the empty set. If the result of the recursive call is the empty set, then the negation is true and we return the singleton set containing the input substitution as a result.

**Dataset:** \{p(a,b), p(a,c), p(b,c), p(c,d)\}

**Call eval:** \(\neg p(Y,d), \{Y \leftarrow b\}\)
**Exit eval:** \(\{Y \leftarrow b\}\) (just one result)

**Call eval:** \(\neg p(Y,d), \{Y \leftarrow c\}\)
**Exit eval:** (no results)
(3) If the expression is a conjunction, we call eval on the first conjunct and the given substitution. We iterate over the list of answers, for each substitution calling eval on the remaining conjuncts and that substitution and return the resulting substitutions.

Dataset: \{p(a,b), p(a,c), p(b,c), p(c,d)\}

Call eval: \(p(a,Y) \land \neg p(Y,d)\), \{
  Call eval: \(p(a,Y)\), \{
    Exit eval: \{Y ← b\}, \{Y ← c\}
  
  Call eval: \(\neg p(Y,d)\), \{Y ← b\}
  Exit eval: \{Y ← b\} (just one result)
  
  Call eval: \(\neg p(Y,d)\), \{Y ← c\}
  Exit eval: (no results)

Exit eval: \{Y ← b\} (just one result)
Evaluation with Variables

Given a query rule, call the procedure `eval` (to be described) on the body and an empty substitution. The result is a list of substitutions that satisfy the body. The value of the rule is obtained by applying the substitutions to the head of the rule.

Dataset: \{p(a,b), p(a,c), p(b,c), p(c,d)\}
Query: `goal(Y) :- p(a,Y) & ~p(Y,d)`

Call eval: `p(a,Y) & ~p(Y,d), {}`
Exit eval: `{Y←b}`

Answer: `{goal(b)}`
Computational Analysis
Worst Case Analysis based on \textit{number of unifications}
\(n\) objects in Herbrand Universe

Assumptions:
- All rules applied
- Subgoals processed left to right

(1) We will first consider analysis with \textit{no indexing}.
(2) Then we will look at analysis with \textit{indexed data}.

\textbf{Optimizations not considered:}
  - Dropping redundant rules or subgoals
  - Reordering of subgoals
  - Caching
Example

\[ \text{goal}(a,c) := p(a,Y) \land p(Y,c) \]

Cost of computing whether \( \text{goal}(a,c) \) is true
Example

\begin{equation}
goal(a,c) :- p(a,Y) \& p(Y,c)\end{equation}

Cost of computing whether \texttt{goal(a,c)} is true

\begin{equation}n^2 + n \times n^2 = n^2 + n^3\end{equation}

Suppose \(n = 3\)

\begin{equation}3^2 + 3 \times 3^2 = 3^2 + 3^3 = 36\end{equation}
Example

\texttt{goal(X,Z) :- p(X,Y) \& p(Y,Z)}

Cost of computing \textit{all} instances of \texttt{goal}
Example

Example

\[ \text{goal}(X,Z) := \ p(X,Y) \ & \ p(Y,Z) \]

Cost of computing \textit{all} instances of \texttt{goal}

\[ n^2 + n^2 \times n^2 = n^2 + n^4 \]

Suppose \( n = 3 \)

\[ 3^2 + 3^2 \times 3^2 = 3^2 + 3^4 = 90 \]
In full indexing, each factoid appears on the list of factoids associated with each constant in that factoid.

Example: \{p(a, b), p(b, c), q(b), q(c)\}

Index on p: \{p(a, b), p(b, c)\}
Index on q: \{q(b), q(c)\}

Index on a: \{p(a, b)\}
Index on b: \{p(a, b), p(b, c), q(b)\}
Index on c: \{p(b, c), q(c)\}

NB: No compound indices (e.g. all factoids with a and b).
Example:

\[
\begin{array}{ccc}
p(a,a) & p(b,a) & p(c,a) \\
p(a,b) & p(b,b) & p(c,b) \\
p(a,c) & p(b,c) & p(c,c) \\
\end{array}
\]

Index on \( p \): \{p(a,a), \ldots, p(c,c)\}
Index on \( a \): \{p(a,a), p(a,b), p(a,c), p(b,a), p(c,a)\}
Index on \( b \): \{p(a,b), p(b,a), p(b,b), p(b,c), p(c,b)\}
Index on \( c \): \{p(a,c), p(b,c), p(c,a), p(c,b), p(c,c)\}
Example with Indexing

Example

\[
\text{goal}(a, c) \ :- \ \text{p}(a, Y) \ & \ \text{p}(Y, c)
\]

Cost of computing \(\text{goal}(a, c)\) \text{ without indexing}

\[
n^2 + n \times n^2 = n^2 + n^3
\]

Suppose \(n = 3\)

\[
3^2 + 3 \times 3^2 = 3^2 + 3^3 = 36
\]

Cost of computing whether \(\text{goal}(a, c)\) \text{ with indexing}

\[
(2n-1) + n \times (2n-1) = 2n^2 + n - 1
\]

Suppose \(n = 3\)

\[
2 \times 3^2 + 3 - 1 = 18 + 2 = 20
\]
Example

\[ \text{goal}(X,Z) := p(X,Y) \land p(Y,Z) \]

Cost of computing \text{goal}(X,Z) \textit{without} indexing

\[ n^2 + n^2 \times n^2 = n^2 + n^4 \]

Suppose \( n = 3 \)

\[ 3^2 + 3^2 \times 3^2 = 3^2 + 3^4 = 90 \]

Cost of computing \textit{all} instances \textit{with} indexing:

\[ n^2 + n^2 \times (2n-1) = n^2 + 2n^3 - n^2 = 2n^3 \]

Suppose \( n = 3 \)

\[ 2 \times 3^3 = 54 \]
\[ p(a,b) \]
\[ p(b,c) \]
\[ p(c,d) \]
Lambda

\[ p(a,b) \]
\[ p(b,c) \]
\[ p(c,d) \]

Query

Pattern: \( \text{goal}(X, Z) \)
Query: \( p(X,Y) \land p(Y,Z) \)

Goal

\[ \text{goal}(a,c) \]
\[ \text{goal}(b,d) \]
Lambda

\[ p(a,b) \\
p(b,c) \\
p(c,d) \]

\[ \text{Pattern} \quad \text{goal}(X,Z) \]
\[ \text{Query} \quad p(X,Y) \land p(Y,Z) \]

\[ \text{goal}(a,c) \]
\[ \text{goal}(b,d) \]
Lambda

Pattern: goal(X, Z)
Query: p(X, Y) & p(Y, Z)

Autorefresh

goal(a, c)
goal(b, d)
\begin{align*}
\text{Goal:} & \quad \text{goal}(X, Z) \\
\text{Query:} & \quad p(X, Y) \land p(Y, Z) \\
\text{Results:} & \quad \text{goal}(a, c), \text{goal}(b, d), \text{goal}(c, e)
\end{align*}
Lambda

Save  Revert  Sort

\begin{align*}
p(a,a) \\
p(a,b) \\
p(a,c) \\
p(b,a) \\
p(b,b) \\
p(b,c) \\
p(c,a) \\
p(c,b) \\
p(c,c)
\end{align*}

Query

Pattern \( \text{goal}(a,c) \)

Query \( p(a,Y) \land p(Y,c) \land \text{false} \)

Show  Next  100 result(s)  Autorefresh

20 unification(s)
Lambda

Save  Revert  Sort

\begin{align*}
p(a,a) \\
p(a,b) \\
p(a,c) \\
p(b,a) \\
p(b,b) \\
p(b,c) \\
p(c,a) \\
p(c,b) \\
p(c,c) \\
\end{align*}

Query

Pattern: goal(X, Z)

Query: \( p(X,Y) \land p(Y,Z) \land \text{false} \)

Show  Next  100  result(s)  Autorefresh

54 unification(s)
**Example**

**Dataset:** \{p(a,b), p(b,c), q(b)\}

**Query:** \texttt{goal(X,Y) :- p(X,Y) \& q(Y)}

**Trace** (for one answer):

- Call: p(X,Y)
- Exit: p(a,b)
- Call: q(b)
- Exit: q(b)
Example

Dataset: \{p(a,b), p(b,c), q(c)\}

Query: goal(X,Y) :- p(X,Y) & q(Y)

Trace (for one answer):
- Call: p(X,Y)  Redo: p(X,Y)
- Exit: p(a,b)  Exit: p(b,c)
- Call: q(b)   Call: q(c)
- Fail: q(b)  Exit: q(c)
Example

Dataset: \{p(a,b), p(b,c), q(b), q(c)\}

Query: goal(X,Y) :- p(X,Y) & q(Y)

Trace (for all answers):

  Call: p(X,Y)         Redo: p(X,Y)         Redo: p(X,Y)
  Exit: p(a,b)         Exit: p(b,c)         Fail: p(X,Y)
  Call: q(b)           Call: q(c)
  Exit: q(b)           Exit: q(c)
  Redo: q(b)           Redo: q(c)
  Fail: q(b)           Fail: q(c)
\[ p(a,b) \]
\[ p(b,c) \]
\[ q(b) \]
Lambda

Save Revert Sort

p(a,a)
p(b,a)
p(c,a)
p(a,b)
p(b,b)
p(c,b)
p(a,c)
p(b,c)
p(c,c)

Query

Pattern  goal(a,c)
Query  p(a,Y) & p(Y,c) & false

Show Next 100 result(s) Autorefresh

20 unification(s)

Specify maximum number of unifications:

100000

Cancel  OK