Title: Inference in First-Order Logic
AIMA: Chapter 9

Introduction to Artificial Intelligence
CSCE 476-876, Spring 2006
URL: www.cse.unl.edu/~choueiry/S06-476-876

Berthe Y. Choueiry (Shu-we-ri)
choueiry@cse.unl.edu, (402)472-5444

Outline

- Reducing first order inference to propositional inference:
  Universal Instantiation, Existential Instantiation,
  Skolemization, Generalized Modus Ponens

- Unification

- Inference mechanisms in First-Order Logic:
  - Forward chaining
  - Backward chaining
  - Resolution (and CNF)
**Universal instantiation** (UI)

Every instantiation of a universally quantified sentence is entailed by it:

\[
\forall v \alpha \\
\text{Subst}(\{v/g\}, \alpha)
\]

for any variable \(v\) and ground term \(g\)

E.g., \(\forall x \text{King}(x) \land \text{Greedy}(x) \Rightarrow \text{Evil}(x)\) yields:

\[
\text{King}(\text{John}) \land \text{Greedy}(\text{John}) \Rightarrow \text{Evil}(\text{John})
\]

\[
\text{King}(\text{Richard}) \land \text{Greedy}(\text{Richard}) \Rightarrow \text{Evil}(\text{Richard})
\]

\[
\text{King}(\text{Father}(\text{John})) \land \text{Greedy}(\text{Father}(\text{John})) \Rightarrow \text{Evil}(\text{Father}(\text{John}))
\]

\[\vdots\]

**Existential instantiation** (EI)

For any sentence \(\alpha\), variable \(v\), and constant symbol \(k\) that does not appear elsewhere in the knowledge base:

\[
\exists v \alpha \\
\text{Subst}(\{v/k\}, \alpha)
\]

E.g., \(\exists x \text{Crown}(x) \land \text{OnHead}(x, \text{John})\) yields

\[
\text{Crown}(C_1) \land \text{OnHead}(C_1, \text{John})
\]

provided \(C_1\) is a new constant symbol, called a **Skolem constant**

Another example: from \(\exists x d(x^y)/dy = x^y\) we obtain

\[
d(e^y)/dy = e^y
\]

provided \(e\) is a new constant symbol
UI and EI

UI can be applied several times to add new sentences; the new KB is logically equivalent to the old.

EI can be applied once to replace the existential sentence; the new KB is not equivalent to the old, but is satisfiable iff the old KB was satisfiable.

Reduction to propositional inference (I)

∀ x King(x) \land Greedy(x) \Rightarrow Evil(x)
King(John)
Greedy(John)
Brother(Richard, John)

Instantiating the universal sentence in all possible ways, we have:

King(John) \land Greedy(John) \Rightarrow Evil(John)
King(Richard) \land Greedy(Richard) \Rightarrow Evil(Richard)
King(John)
Greedy(John)
Brother(Richard, John)

The new KB is propositionalized: proposition symbols are:
King(John), Greedy(John), Evil(John), King(Richard) etc.
Reduction to propositional inference (II)

- Claim: a ground sentence* is entailed by new KB iff entailed by original KB
- Claim: every FOL KB can be propositionalized so as to preserve entailment
- Idea: propositionalize KB and query, apply resolution, return result
- Problem: with function symbols, there are infinitely many ground terms, e.g., Father(Father(Father(John)))

Reduction to propositional inference (III)

- Theorem: Herbrand (1930). If a sentence $\alpha$ is entailed by an FOL KB, it is entailed by a finite subset of the propositional KB
- Idea: For $n = 0$ to $\infty$ do
  create a propositional KB by instantiating with depth-$n$ terms
  see if $\alpha$ is entailed by this KB
- Problem: works if $\alpha$ is entailed, loops if $\alpha$ is not entailed
- Theorem: Turing (1936), Church (1936), entailment in FOL is semidecidable
Problems with propositionalization

Propositionalization generates lots of irrelevant sentences. E.g., from
\[ \forall x King(x) \land Greedy(x) \Rightarrow Evil(x) \]
King(John)
\[ \forall y Greedy(y) \]
Brother(Richard, John)
it seems obvious that Evil(John), but propositionalization produces lots of facts such as Greedy(Richard) that are irrelevant
With \( p \) \( k \)-ary predicates and \( n \) constants, there are \( p \cdot n^k \) instantiations!

Unification

We can get the inference immediately if we can find a substitution \( \theta \) such that \( King(x) \) and \( Greedy(x) \) match \( King(John) \) and \( Greedy(y) \)
\[ \theta = \{ x/John, y/John \} \] works
Unify(\( \alpha, \beta \)) = \( \theta \) if \( \alpha \theta = \beta \theta \)

<table>
<thead>
<tr>
<th>( p )</th>
<th>( q )</th>
<th>( \theta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knows(John, x)</td>
<td>Knows(John, Jane)</td>
<td>{x/Jane}</td>
</tr>
<tr>
<td>Knows(John, x)</td>
<td>Knows(y, OJ)</td>
<td>{x/OJ, y/John}</td>
</tr>
<tr>
<td>Knows(John, x)</td>
<td>Knows(y, Mother(y))</td>
<td>{y/John, x/Mother(John)}</td>
</tr>
<tr>
<td>Knows(John, x)</td>
<td>Knows(x, OJ)</td>
<td>fail</td>
</tr>
</tbody>
</table>

Standardizing apart eliminates overlap of variables, e.g.,
Knows(\( z_{17} \), OJ)
**Generalized Modus Ponens (GMP)**

\[
p_1', p_2', \ldots, p_n', (p_1 \land p_2 \land \ldots \land p_n \Rightarrow q) \quad \therefore q^\theta
\]

where \( p_i\theta = p_i \) for all \( i \)

- \( p_1' \) is \( \text{King}(\text{John}) \)
- \( p_1 \) is \( \text{King}(x) \)
- \( p_2' \) is \( \text{Greedy}(y) \)
- \( p_2 \) is \( \text{Greedy}(x) \)
- \( \theta \) is \( \{x/\text{John}, y/\text{John}\} \)
- \( q \) is \( \text{Evil}(x) \)
- \( q^\theta \) is \( \text{Evil}(\text{John}) \)

GMP used with KB of **definite clauses** (*exactly* one positive literal)

All variables assumed universally quantified

---

**Example knowledge base**

The law says that it is a crime for an American to sell weapons to hostile nations. The country Nono, an enemy of America, has some missiles, and all of its missiles were sold to it by Colonel West, who is American.

Prove that Col. West is a criminal

\[\ldots \text{it is a crime for an American to sell weapons to hostile nations:} \]
\[\text{American}(x) \land \text{Weapon}(y) \land \text{Sells}(x, y, z) \land \text{Hostile}(z) \Rightarrow \text{Criminal}(x)\]
Example of KB (2)

Nono ... has some missiles, i.e., \( \exists x \, \text{Owns}(\text{Nono}, x) \land \text{Missile}(x) \): \( \text{Owns}(\text{Nono}, \text{M}_1) \) and \( \text{Missile}(\text{M}_1) \)

... all of its missiles were sold to it by Colonel West
\( \forall x \, \text{Missile}(x) \land \text{Owns}(\text{Nono}, x) \Rightarrow \text{Sells}(\text{West}, x, \text{Nono}) \)

Missiles are weapons:
\( \text{Missile}(x) \Rightarrow \text{Weapon}(x) \)

Example of KB (3)

An enemy of America counts as “hostile”:
\( \text{Enemy}(x, \text{America}) \Rightarrow \text{Hostile}(x) \)

West, who is American ...
\( \text{American}(\text{West}) \)

The country Nono, an enemy of America ...
\( \text{Enemy}(\text{Nono}, \text{America}) \)
Forward chaining algorithm

< FOL-FC-Ask, Figure 9.3 page 282 >

Forward chaining proof

```
Criminal(West)

Weapon(M₁)  Sells(West,M₁,Nono)  Hostile(Nono)

American(West)  Missile(M₁)  Owns(Nono,M₁)  Enemy(Nono,America)
```
Properties of forward chaining

- Sound and complete for first-order definite clauses (proof similar to propositional proof)
- Datalog = first-order definite clauses + no functions (e.g., crime KB)
  FC terminates for Datalog in poly iterations: at most $p \cdot n^k$
literals
- May not terminate in general if $\alpha$ is not entailed
- This is unavoidable: entailment with definite clauses is semidecidable

Efficiency of forward chaining

- Simple observation: no need to match a rule on iteration $k$
  if a premise wasn’t added on iteration $k - 1$
  $\Rightarrow$ match each rule whose premise contains a newly added
  literal
- Matching itself can be expensive
- Database indexing allows $O(1)$ retrieval of known facts
  e.g., query $\text{Missile}(x)$ retrieves $\text{Missile}(M_1)$
- Matching conjunctive premises against known facts is NP-hard
- Forward chaining is widely used in deductive databases
Backward chaining algorithm

<FO-L-BC-Ask, Figure 9.6 page 288>
Properties of backward chaining

- Depth-first recursive proof search: space is linear in size of proof
- Incomplete due to infinite loops
  ⇒ fix by checking current goal against every goal on stack
- Inefficient due to repeated subgoals (both success and failure)
  ⇒ fix using caching of previous results (extra space!)
- Widely used (without improvements!) for logic programming

Resolution: brief summary

Full first-order version:

\[
\frac{l_1 \lor \ldots \lor l_i, \quad m_1 \lor \ldots \lor m_n}{(l_1 \lor \ldots \lor l_{i-1} \lor l_{i+1} \lor \ldots \lor l_k \lor m_1 \lor \ldots \lor m_{j-1} \lor m_{j+1} \lor \ldots \lor m_n)\theta}
\]

where Unify\((l_i, \neg m_j) = \theta\).

For example,

\[
\frac{\neg \text{Rich}(x) \lor \text{Unhappy}(x) \quad \text{Rich}(\text{Ken})}{\text{Unhappy}(\text{Ken})}
\]

with \(\theta = \{x/\text{Ken}\}\)

Apply resolution steps to CNF\((KB \land \neg \alpha)\); complete for FOL
Conversion to CNF (I)

Everyone who loves all animals is loved by someone:
\[ \forall x [\forall y \text{Animal}(y) \Rightarrow \text{Loves}(x, y)] \Rightarrow [\exists y \text{Loves}(y, x)] \]

1. Eliminate biconditionals and implications

\[ \forall x [\neg \forall y \neg \text{Animal}(y) \lor \text{Loves}(x, y)] \lor [\exists y \text{Loves}(y, x)] \]

2. Move \( \neg \) inwards: \( \neg \forall x, p \equiv \exists x \neg p \)

\[ \forall x [\exists y (\neg \text{Animal}(y) \lor \text{Loves}(x, y))] \lor [\exists y \text{Loves}(y, x)] \]

\[ \forall x [\exists y \neg \text{Animal}(y) \land \neg \text{Loves}(x, y)] \lor [\exists y \text{Loves}(y, x)] \]

\[ \forall x [\exists y \text{Animal}(y) \land \neg \text{Loves}(x, y)] \lor [\exists y \text{Loves}(y, x)] \]

Conversion to CNF (II)

3. Standardize variables: each quantifier should use a different one

\[ \forall x [\exists y \text{Animal}(y) \land \neg \text{Loves}(x, y)] \lor [\exists z \text{Loves}(z, x)] \]

4. Skolemize: a more general form of existential instantiation. Each existential variable is replaced by a Skolem function of the enclosing universally quantified variables:

\[ \forall x [\text{Animal}(F(x)) \land \neg \text{Loves}(x, F(x))] \lor \text{Loves}(G(x), x) \]

5. Drop universal quantifiers:

\[ [\text{Animal}(F(x)) \land \neg \text{Loves}(x, F(x))] \lor \text{Loves}(G(x), x) \]

6. Distribute \( \land \) over \( \lor \):

\[ [\text{Animal}(F(x)) \lor \text{Loves}(G(x), x)] \land [\neg \text{Loves}(x, F(x)) \lor \text{Loves}(G(x), x)] \]
Resolution proof: definite clauses

\[ \neg \text{American}(x) \lor \neg \text{Weapon}(y) \lor \neg \text{Sells}(x,y,z) \lor \neg \text{Hostile}(z) \lor \text{Criminal}(x) \]

\[ \neg \text{Criminal}(\text{West}) \]

\[ \neg \text{Enemy}(\text{Nono}, \text{America}) \]

\[ \neg \text{American}(\text{West}) \lor \neg \text{Weapon}(y) \lor \neg \text{Sells}(\text{West},y,z) \lor \neg \text{Hostile}(z) \]

\[ \neg \text{Missile}(x) \lor \text{Weapon}(x) \]

\[ \neg \text{Missile}(x) \lor \neg \text{Owns}(\text{Nono},x) \lor \text{Sells}(\text{West},x,\text{Nono}) \]

\[ \neg \text{Sells}(\text{West},x,\text{Nono}) \lor \neg \text{Hostile}(\text{West}) \]

\[ \neg \text{Missile}(\text{M}_1) \lor \neg \text{Owns}(\text{Nono}, \text{M}_1) \lor \neg \text{Hostile}(\text{Nono}) \]

\[ \neg \text{Missile}(\text{M}_1) \lor \neg \text{Owns}(\text{Nono}, \text{M}_1) \lor \neg \text{Hostile}(\text{Nono}) \]

\[ \neg \text{Missile}(\text{M}_1) \lor \neg \text{Owns}(\text{Nono}, \text{M}_1) \lor \neg \text{Hostile}(\text{Nono}) \]

\[ \neg \text{Enemy}(\text{x,America}) \lor \text{Hostile}(\text{x}) \]

\[ \neg \text{Hostile}(\text{Nono}) \]

\[ \neg \text{Enemy}(\text{Nono,America}) \]