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:

\[ \forall v \alpha \]

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:

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

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

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

\[ \vdots \]
**Existential instantiation (EI)**

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

$$\exists v \alpha$$

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 \textit{add} new sentences; the new KB is logically equivalent to the old.

EI can be applied once to \textit{replace} the existential sentence; the new KB is \textit{not} equivalent to the old, but is satisfiable iff the old KB was satisfiable.
Reduction to propositional inference (I)

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

\text{King}(John)

\text{Greedy}(John)

\text{Brother}(Richard, John)

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

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

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

\text{King}(John)

\text{Greedy}(John)

\text{Brother}(Richard, John)

The new KB is propositionalized: proposition symbols are:

\text{King}(John), \text{Greedy}(John), \text{Evil}(John), \text{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 negated query, put in CNF, apply resolution, derive contradiction

- Problem: with function symbols, there are infinitely many ground terms, e.g., \( \text{Father}(\text{Father}(\text{Father}(\text{John}))) \)
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

\[ \forall x \, \text{King}(x) \land \text{Greedy}(x) \Rightarrow \text{Evil}(x) \]
\[ \forall y \, \text{Greedy}(y) \]
\[ \text{King}(\text{John}) \]
\[ \text{Brother}(\text{Richard}, \text{John}) \]

it seems obvious that \( \text{Evil}(\text{John}) \), but propositionalization produces lots of facts such as \( \text{Greedy}(\text{Richard}) \) 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 $\text{King}(x)$ and $\text{Greedy}(x)$ match $\text{King}(\text{John})$ and $\text{Greedy}(y)$.

$$\theta = \{x/\text{John}, y/\text{John}\}$$ 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>$\text{Knows}(\text{John}, x)$</td>
<td>$\text{Knows}(\text{John}, \text{Jane})$</td>
<td>${x/\text{Jane}}$</td>
</tr>
<tr>
<td>$\text{Knows}(\text{John}, x)$</td>
<td>$\text{Knows}(y, \text{OJ})$</td>
<td>${x/\text{OJ}, y/\text{John}}$</td>
</tr>
<tr>
<td>$\text{Knows}(\text{John}, x)$</td>
<td>$\text{Knows}(y, \text{Mother}(y))$</td>
<td>${y/\text{John}, x/\text{Mother}(\text{John})}$</td>
</tr>
<tr>
<td>$\text{Knows}(\text{John}, x)$</td>
<td>$\text{Knows}(x, \text{OJ})$</td>
<td>fail</td>
</tr>
</tbody>
</table>

Standardizing apart eliminates overlap of variables, e.g.,

$\text{Knows}(z_{17}, \text{OJ})$
Generalized Modus Ponens (GMP)

\[
p_1', p_2', \ldots, p_n', (p_1 \land p_2 \land \ldots \land p_n \Rightarrow q) \quad \frac{q_{\theta}}{q_{\theta}}
\]

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

\( p_1' \) is \( King(John) \) \quad \( p_1 \) is \( King(x) \)
\( p_2' \) is \( Greedy(y) \) \quad \( p_2 \) is \( Greedy(x) \)
\( \theta \) is \( \{x/John, y/John\} \) \quad \( q \) is \( Evil(x) \)
\( q_{\theta} \) is \( Evil(John) \)

GMP used with KB of **definite clauses** (**exactly** one positive literal)
All variables assumed universally quantified
Applying GMP: on an (artificial) example

Consider the (bogus) KB with definite clauses:

1. $\forall x_1 P(x_1, A)$
2. $\forall x_2 Q(B, x_2)$
3. $\forall x_3, x_4, x_5 P(C, x_3) \land Q(x_4, x_5) \rightarrow R(C, x_3, x_4, x_5)$

All variables are universally quantified:

1. $P(x_1, A)$
2. $Q(B, x_2)$
3. $P(C, x_3) \land Q(x_4, x_5) \rightarrow R(C, x_3, x_4, x_5)$

Applying GMP with $\theta=\{x_1/C, x_3/A, x_4/B, x_2/x_5\}$, we add to KB:

$$R(C, A, B, x_2)$$
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

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

Nono . . . has some missiles, i.e., $\exists x \; Owns(Nono, x) \land Missile(x)$: $Owns(Nono, M_1)$ and $Missile(M_1)$

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

Missiles are weapons:
$Missile(x) \Rightarrow Weapon(x)$
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 . . .
\[ \text{Enemy}(\text{Nono}, \text{America}) \]
The considered KB becomes:

1. \( \text{American}(x) \land \text{Weapon}(y) \land \text{Sells}(x, y, z) \land \text{Hostile}(z) \Rightarrow \text{Criminal}(x) \)

2. \( \text{Owns}(\text{Nono}, M_1) \)

3. \( \text{Missile}(M_1) \)

4. \( \text{Missile}(x) \land \text{Owns}(\text{Nono}, x) \Rightarrow \text{Sells}(\text{West}, x, \text{Nono}) \)

5. \( \text{Missile}(x) \Rightarrow \text{Weapon}(x) \)

6. \( \text{Enemy}(x, \text{America}) \Rightarrow \text{Hostile}(x) \)

7. \( \text{American}(\text{West}) \)

8. \( \text{Enemy}(\text{Nono}, \text{America}) \)

We can apply GMP by chaining the rules forward or backward on this KB because it is a Horn theory (definite clauses)
Forward chaining algorithm

<FOL-FC-Ask, Figure 9.3 page 288>
Forward chaining proof

1. Criminal(West)
2. American(West)
3. Missile(M1)
4. Owns(Nono,M1)
5. Sells(West,M1,Nono)
6. Hostile(Nono)
7. 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

<FOL-BC-Ask, Figure 9.6 page 293>
Backward chaining example

- Criminal(West)
  - American(West)
    - Weapon(y)
      - Missile(y) {y/M1}
    - Sells(West,M1,z) {z/Nono}
  - Owns(Nono,M1) { }
  - Enemy(Nono,America) { }
  - Hostile(Nono) { }
Properties of backward chaining

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

Full first-order version:

\[
\begin{array}{c}
l_1 \lor \cdots \lor l_k, \quad m_1 \lor \cdots \lor m_n \\
(l_1 \lor \cdots \lor l_{i-1} \lor l_{i+1} \lor \cdots \lor l_k \lor m_1 \lor \cdots \lor m_{j-1} \lor m_{j+1} \lor \cdots \lor m_n)\theta
\end{array}
\]

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

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 \(\text{CNF}(KB \land \neg \alpha)\); complete for FOL
Conversion to CNF (I)

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

1. Eliminate biconditionals and implications

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

2. Move \( \neg \) inwards:

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

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

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

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

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

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

$$\forall x [Animal(F(x)) \land \neg Loves(x, F(x))] \lor Loves(G(x), x)$$

5. Drop universal quantifiers:

$$[Animal(F(x)) \land \neg Loves(x, F(x))] \lor Loves(G(x), x)$$

6. Distribute $\land$ over $\lor$:

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

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

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

\[ \text{American}(\text{West}) \]

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

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

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

\[ \text{Missile}(M) \]

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

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

\[ \neg \text{Sells}(\text{West},M,z) \lor \neg \text{Hostile}(z) \]

\[ \text{Missile}(M) \]

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

\[ \text{Owns}(\text{Nono},M) \]

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

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

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

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

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