Inference Strategy: Backward Chaining

Idea:
- Check whether a particular fact \( q \) is true.

Backward Chaining:
Given a fact \( q \) to be "proven",
1. See if \( q \) is already in the KB. If so, return TRUE.
2. Find all implications, \( I \), whose conclusion "matches" \( q \).
3. Recursively establish the premises of all \( i \) in \( I \) via backward chaining.

Avoids inferring unrelated facts.

Example: Backward Chaining

Knowledge-base describing when the car should brake?
- \((\text{PersonInFrontOfCar} \implies \text{Brake})\)
- \((\text{YellowLight} \land \text{Policeman} \land (\neg \text{Slippery}) \implies \text{Brake})\)
- \((\text{Policeman} \implies \text{Brake})\)
- \((\text{Snow} \implies \text{Slippery})\)
- \((\text{Slippery} \implies \neg \text{Dry})\)
- \((\text{RedLight} \implies \text{Brake})\)
- \((\text{Winter} \implies \text{Snow})\)

Observation from sensors:
- \((\text{YellowLight} \land \neg \text{RedLight} \land \neg \text{Snow} \land \text{Dry} \land \text{Policeman} \land \neg \text{PersonInFrontOfCar})\)

Should the agent brake (i.e., can "brake" be inferred)?
- Goal: Brake
  - Modus Ponens (brake): \( \text{PersonInFrontOfCar} \)
  - Failure: \( \text{PersonInFrontOfCar} \)

- Backtracking
  - Goal: Brake
    - Modus Ponens (brake): \( \text{YellowLight} \land \text{Policeman} \land (\neg \text{Slippery}) \)
    - Known (YellowLight): \( \text{Policeman} \land (\neg \text{Slippery}) \)
    - Modus Ponens (Policeman): \( \text{Policecar} \land (\neg \text{Slippery}) \)
    - Known (Policecar): \( \neg \text{Slippery} \)
    - Modus Tollens (\( \neg \text{Slippery} \)): \( \text{Dry} \)
    - Known (Dry)

First-Order Logic

Idea:
- Don’t treat propositions as “atomic” entities.

First-Order Logic:
- Objects: \( \text{cs472}, \text{fred}, \text{oh155}, \text{emptylist} \) …
- Relations/Predicates: \( \text{is.Man(fred)}, \text{Located(cs472, oh155)} \) …
  - Note: Relations typically correspond to verbs
- Functions: \( \text{Pair(search,Pair(learning,Pair(kbsystems, emptylist)))} \)
- Connectives: \( \land, \lor, \neg, \Rightarrow, \Leftrightarrow \)
- Quantifiers:
  - Universal: \( \forall \ x : ( \text{is.Man}(x) \Rightarrow \text{is.Mortal}(x) ) \)
  - Existential: \( \exists \ y : ( \text{is.Father}(y, \text{fred}) ) \)

Example: Representing Facts in First-Order Logic

1. Lucy* is a professor
2. All professors are people.
3. Fuchs is the dean.
4. Deans are professors.
5. All professors consider the dean a friend or don’t know him.
6. Everyone is a friend of someone.
7. People only criticize people that are not their friends.
8. Lucy criticized Fuchs.

* Name changed for privacy reasons.

Example: Proof

Knowledge base:
- \( \text{is.prof(lucy)} \)
- \( \forall \ x ( \text{is.prof}(x) \rightarrow \text{is.person}(x) ) \)
- \( \text{is.dean(fuchs)} \)
- \( \forall \ x ( \text{is.dean}(x) \rightarrow \text{is.prof}(x)) \)
- \( \forall \ x ( \forall \ y ( \text{is.prof}(x) \land \text{is.dean}(y) \rightarrow \text{is.friend-off}(y, x) \lor \neg \text{knows}(x, y) )) \)
- \( \forall \ x ( \exists \ y ( \text{is.friend-of}(y, x) )) \)
- \( \forall \ x ( \forall \ y ( \text{is.person}(x) \land \text{is.person}(y) \land \text{criticize}(x, y) \rightarrow \neg \text{is.friend-of}(y, x))) \)
- \( \text{criticize(lucy,fuchs)} \)

Question: Is Fuchs no friend of Lucy?
- \( \neg \text{is.friend-of(fuchs, lucy)} \)

Knowledge Engineering

1. Identify the task.
2. Assemble the relevant knowledge.
3. Decide on a vocabulary of predicates, functions, and constants.
4. Encode general knowledge about the domain.
5. Encode a description of the specific problem instance.
6. Pose queries to the inference procedure and get answers.
7. Debug the knowledge base.
Inference Procedures: Theoretical Results

- There exist complete and sound proof procedures for propositional and FOL.
  - Propositional logic
    - Use the definition of entailment directly. Proof procedure is exponential in n, the number of symbols.
    - In practice, can be much faster...
    - Polynomial-time inference procedure exists when KB is expressed as Horn clauses: \( P_1 \land P_2 \land \ldots \land P_n \Rightarrow Q \)
      where the \( P_i \) and \( Q \) are non-negated atoms.
  - First-Order logic
    - Gödel’s completeness theorem showed that a proof procedure exists...
    - But none was demonstrated until Robinson’s 1965 resolution algorithm.
    - Entailment in first-order logic is semidecidable.

Resolution Rule of Inference

General Rule:

Assume: \( E \lor E_{12} \lor \ldots \lor E_{1k} \)
and \( \neg E \lor E_{22} \lor \ldots \lor E_{2l} \)

Then: \( E_{12} \lor \ldots \lor E_{1k} \lor E_{22} \lor \ldots \lor E_{2l} \)

Note: \( E_i \) can be negated.

Example:

Assume: \( E_1 \lor E_2 \) playing tennis or raining
and \( \neg E_2 \lor E_3 \) not raining or working

Then: \( E_1 \lor E_3 \) playing tennis or working

Algorithm: Resolution Proof

- Negate the theorem to be proved, and add the result to the knowledge base.
- Bring knowledge base into conjunctive normal form (CNF)
  - CNF: conjunctions of disjunctions
  - Each disjunction is called a clause.
- Until there is no resolvable pair of clauses,
  - Find resolvable clauses and resolve them.
  - Add the results of resolution to the knowledge base.
  - If NIL (empty clause) is produced, stop and report that the (original) theorem is true.
- Report that the (original) theorem is false.

Resolution Example: Propositional Logic

- To prove: \( \neg P \)
- Transform Knowledge Base into CNF

<table>
<thead>
<tr>
<th>Regular</th>
<th>CNF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentence 1: ( P \lor Q )</td>
<td>( \neg P \lor Q )</td>
</tr>
<tr>
<td>Sentence 2: ( Q \lor R )</td>
<td>( \neg Q \lor R )</td>
</tr>
<tr>
<td>Sentence 3: ( \neg R )</td>
<td>( \neg R )</td>
</tr>
</tbody>
</table>

- Proof
  1. \( \neg P \lor Q \) Sentence 1
  2. \( \neg Q \lor R \) Sentence 2
  3. \( \neg R \) Sentence 3
  4. Assume opposite
  5. \( Q \) Resolve 4 and 1
  6. \( R \) Resolve 5 and 2
  7. nil Resolve 6 with 3

Resolution Example: FOL

Example: Prove bird (tweety)

Axioms: Regular + CNF

1. \( \forall x: \text{feathers}(x) \rightarrow \text{bird}(x) \)  \( \neg \text{feathers}(x) \lor \text{bird}(x) \)
2. \( \text{feathers}(\text{tweety}) \) \( \text{feathers}(\text{tweety}) \)
3. \( \neg \text{bird}(\text{tweety}) \) \( \neg \text{bird}(\text{tweety}) \)
4. Resolution Proof

1. Resolve 3 and 1, specializing (i.e. “unifying”) tweety for x.
   Add \( \neg \text{feathers}(\text{tweety}) \)
2. Resolve 4 and 2. Add NIL.

Resolution Theorem Proving

Properties of Resolution Theorem Proving:
- sound (for propositional and FOL)
- (refutation) complete (for propositional and FOL)

Procedure may seem cumbersome but note that can be easily automated. Just “smash” clauses until empty clause or no more new clauses.
**Unification**

**Unify procedure:** Unify(P,Q) takes two atomic (i.e. single predicates) sentences P and Q and returns a substitution that makes P and Q identical.

**Rules for substitutions:**
- Can replace a variable by a constant.
- Can replace a variable by a variable.
- Can replace a variable by a function expression, as long as the function expression does not contain the variable.

**Unifier:** a substitution that makes two clauses resolvable. 

\[ v_1/C; v_2/v_3; v_4/f(\ldots) \]

**Unification - Purpose**

**Given:**
\[ \neg \text{Knows}(\text{John}, x) \lor \text{Hates}(\text{John}, x) \]
\[ \text{Knows}(\text{John}, \text{Jim}) \]

**Derive:**
\[ \text{Hates}(\text{John}, \text{Jim}) \]

**Unification:**
unify(\text{Knows}(\text{John}, x), \text{Knows}(\text{John}, \text{Jim})) = \{x/\text{Jim}\}

Need unifier \{x/\text{Jim}\} for resolution to work.

**Add to knowledge base:**
\[ \neg \text{Knows}(\text{John}, \text{Jim}) \lor \text{Hates}(\text{John}, \text{Jim}) \]

---

**Unification (example)**

**Who does John hate?** 
\[ \exists x \text{Knows}(\text{John}, x) \]

**Knowledge base (in clause form):**
1. \( \neg \text{Knows}(\text{John}, x) \lor \text{Hates}(\text{John}, x) \)
2. \text{Knows}(\text{John}, \text{Jim})
3. \text{Knows}(\text{John}, \text{Leo})
4. \text{Knows}(z, \text{Mother}(z))
5. \( \neg \text{Hates}(\text{John}, x) \) (since \( \exists x \text{Knows}(\text{John}, x) \land \forall x \neg \text{Hates}(\text{John}, x) \))

**Resolution with 5 and 1:**
unify(\text{Hates}(\text{John}, x), \text{Hates}(\text{John}, v)) = \{x/v\}

**Add to knowledge base:**
\[ \neg \text{Knows}(\text{John}, v) \]

**Resolution with 6 and 2:**
unify(\text{Knows}(\text{John}, v), \text{Knows}(\text{John}, \text{Jim})) = \{v/\text{Jim}\}

**or Resolution with 6 and 3:**
unify(\text{Knows}(\text{John}, v), \text{Knows}(y, \text{Leo})) = \{y/\text{John}, v/\text{Leo}\}

**or Resolution with 6 and 4:**
unify(\text{Knows}(\text{John}, v), \text{Knows}(z, \text{Mother}(z))) = \{z/\text{John}, v/\text{Mother}(z)\}

**Answers:**
1. \text{Hates}(\text{John}, x) with \{x/v, v/\text{Jim}\} (i.e. John hates Jim)
2. \text{Hates}(\text{John}, x) with \{x/v, y/\text{John}, v/\text{Leo}\} (i.e. John hates Leo)
3. \text{Hates}(\text{John}, x) with \{x/v, v/\text{Mother}(z), z/\text{John}\} (i.e. John hates his mother)

---

**Most General Unifier**

In cases where there is more than one substitution choose the one that makes the least commitment (most general) about the bindings.

UNIFY
\[ \text{Knows}(\text{John}, x), \text{Knows}(v, z) \]

= \{y/\text{John}, x/\text{z}\}

or \{y/\text{John}, x/\text{z}, z/\text{Freda}\}

or \{y/\text{John}, x/\text{John}, z/\text{John}\}

...

See R&N for general unification algorithm. \(O(n^2)\) with Refutation

---

**Converting More Complicated Sentences to CNF**

**Sentence:**
\[ \forall x : \text{brick}(x) \rightarrow ((\exists y : \text{on}(x, y) \land \neg \text{pyramid}(y)) \land \neg \exists y : \text{on}(x, y) \land \text{on}(y, x)) \land (\forall y : \neg \text{brick}(y) \rightarrow \neg \text{equal}(x, y)) \]

**CNF:**
\[ \neg \text{brick}(x) \lor \text{on}(x, \text{support}(x)) \]
\[ \neg \text{brick}(y) \lor \neg \text{pyramid}(\text{support}(y)) \]
\[ \neg \text{brick}(z) \lor \neg \text{brick}(u) \lor \neg \text{on}(y, u) \]
\[ \neg \text{brick}(v) \lor \text{brick}(z) \lor \neg \text{equal}(v, z) \]

---

**1. Eliminate Implications**

**Substitute:** \( \neg E_1 \lor E_2 \) for \( E_1 \rightarrow E_2 \)

\[ \forall x : \neg \text{brick}(x) \lor ((\exists y : \text{on}(x, y) \land \neg \text{pyramid}(y)) \land (\neg \exists y : \text{on}(x, y) \land \text{on}(y, x)) \land (\forall y : \neg \text{brick}(y) \lor \neg \text{equal}(x, y)) \]

\[ \forall x : \neg \text{brick}(x) \lor ((\exists y : \text{on}(x, y) \land \neg \text{pyramid}(y)) \land (\neg \exists y : \text{on}(x, y) \land \text{on}(y, x)) \land (\forall y : \neg \text{brick}(y) \lor \neg \text{equal}(x, y)) \]
2. Move negations down to the atomic formulas

Equivalence Transformations:

\[ \neg(E_1 \land E_2) \iff (\neg E_1) \lor (\neg E_2) \]
\[ \neg(E_1 \lor E_2) \iff (\neg E_1) \land (\neg E_2) \]
\[ (\neg E_1) \iff E_1 \]
\[ \forall x : E_1(x) \iff \exists x : \neg E_1(x) \]
\[ \exists x : E_1(x) \iff \forall x : \neg E_1(x) \]

Result:

\[ \forall x : \neg \text{brick}(x) \lor \]
\[ ((\exists y : \text{on}(x,y) \land \neg \text{pyramid}(y))) \]
\[ \land (\neg \exists y : \text{on}(x,y) \land \text{on}(y,x)) \]
\[ \land (\forall y : \neg (\neg \text{brick}(y)) \lor \neg \text{equal}(x,y)) \]

3. Eliminate Existential Quantifiers:

Skolemization

Harder cases:

\[ \forall x : \exists y : \text{father}(y,x) \text{ becomes } \forall x : \text{father}(S1(x), x) \]

There is one argument for each universally quantified variable whose scope contains the Skolem function.

Easy case:

\[ \exists x : \text{President}(x) \text{ becomes } \text{President}(S2) \]
\[ \forall x : \neg \text{brick}(x) \lor ((\exists y : \text{on}(x,y) \land \neg \text{pyramid}(y)) \lor \ldots \]

4. Rename variables as necessary

We want no two variables of the same name.

\[ \forall x : \neg \text{brick}(x) \lor \]
\[ (\exists y : \text{on}(x,y) \land \neg \text{pyramid}(y)) \]
\[ \land (\neg \exists y : \text{on}(x,y) \land \text{on}(y,x)) \]
\[ \land (\forall y : \neg \text{brick}(y) \lor \neg \text{equal}(x,y)) \]

\[ \forall x : \neg \text{brick}(x) \lor \]
\[ (\exists y : \text{on}(x,y) \land \neg \text{pyramid}(y)) \]
\[ \land (\neg \exists y : \text{on}(x,y) \land \text{on}(y,x)) \]
\[ \land (\forall z : \text{brick}(z) \lor \neg \text{equal}(x,z)) \]

5. Move the universal quantifiers to the left

This works because each quantifier uses a unique variable name.

\[ \forall x : \neg \text{brick}(x) \lor (\exists y : \text{on}(x,y) \land \neg \text{pyramid}(y)) \]
\[ \land (\forall z : \text{brick}(z) \lor \neg \text{equal}(x,z)) \]

\[ \forall x : \neg \text{brick}(x) \lor (\exists y : \text{on}(x,y) \land \neg \text{pyramid}(y)) \]
\[ \land (\forall z : \text{brick}(z) \lor \neg \text{equal}(x,z)) \]

\[ \forall x : \neg \text{brick}(x) \lor (\exists y : \text{on}(x,y) \land \neg \text{pyramid}(y)) \]
\[ \land (\forall z : \text{brick}(z) \lor \neg \text{equal}(x,z)) \]

6. Move disjunctions down to the literals

\[ E_1 \lor (E_2 \land E_3) \iff (E_1 \lor E_2) \land (E_1 \lor E_3) \]

\[ \forall x \forall y \forall z : (\neg \text{brick}(x) \lor (\exists y : \text{on}(x,y) \land \neg \text{pyramid}(y))) \]
\[ \land (\neg \exists y : \text{on}(x,y) \land \text{on}(y,x)) \]
\[ \land (\neg \text{brick}(z) \lor \neg \text{equal}(x,z)) \]

\[ \forall x \forall y \forall z : (\neg \text{brick}(x) \lor (\exists y : \text{on}(x,y) \land \neg \text{pyramid}(y))) \]
\[ \land (\neg \exists y : \text{on}(x,y) \land \text{on}(y,x)) \]
\[ \land (\neg \text{brick}(z) \lor \neg \text{equal}(x,z)) \]

7. Eliminate the conjunctions

\[ \forall x \forall y \forall z : (\neg \text{brick}(x) \lor \text{on}(x,S1(x))) \]
\[ \land (\neg \text{brick}(z) \lor \neg \text{pyramid}(S1(x))) \]
\[ \land (\neg \text{brick}(x) \lor \neg \text{on}(x,y) \lor \neg \text{on}(y,x)) \]
\[ \land (\neg \text{brick}(x) \lor \text{brick}(z) \lor \neg \text{equal}(x,z)) \]

\[ \forall x : \neg \text{brick}(x) \lor \text{on}(x,S1(x)) \]
\[ \forall x : \neg \text{brick}(x) \lor \neg \text{pyramid}(S1(x)) \]
\[ \forall x : \neg \text{brick}(x) \lor \neg \text{on}(x,y) \lor \neg \text{on}(y,x) \]
\[ \forall x : \neg \text{brick}(x) \lor \text{brick}(z) \lor \neg \text{equal}(x,z) \]
Algorithm: Putting Axioms into Clausal Form

1. Eliminate the implications.
2. Move the negations down to the atomic formulas.
3. Eliminate the existential quantifiers.
4. Rename the variables, if necessary.
5. Move the universal quantifiers to the left.
6. Move the disjunctions down to the literals.
7. Eliminate the conjunctions.
8. Rename the variables, if necessary.
9. Eliminate the universal quantifiers.

Resolution Proofs as Search

- **Search Problem**
  - States: Content of knowledge base in CNF
  - Initial state: Knowledge base with negated theorem to prove
  - Successor function: Resolution inference rule with unify
  - Goal test: Does knowledge base contain the empty clause ‘nil’

- **Search Algorithm**
  - Depth first search (used in PROLOG)
    - Note: Possibly infinite state space
    - Example:
      - `IsPerson(Fred)`
      - `IsPerson(y) IsPerson(mother(y))`
      - Goal: `∃x: IsPerson(x)`
      - Answers: `{x/Fred}` and `{x/mother(Fred)}` and `{x/mother(mother(Fred))}` and …

Example

Jack owns a dog.

Every dog owner is an animal lover.

No animal lover kills an animal.

Either Jack or Curiosity killed the cat, who is named Tuna.

Did Curiosity kill the cat?
**Conjunctive Normal Form**

\[
\begin{align*}
& \text{Dog}(D) \\
& \text{Owns}(\text{Jack}, D) \\
& \neg \text{Dog}(y) \lor \neg \text{Owns}(x, y) \lor \text{AnimalLover}(x) \\
& \neg \text{AnimalLover}(w) \lor \neg \text{Animal}(y) \lor \neg \text{Kills}(w, y) \\
& \text{Kills}(\text{Jack}, \text{Tuna}) \lor \text{Kills}(\text{Curiosity}, \text{Tuna}) \\
& \text{Cat}(\text{Tuna}) \\
& \neg \text{Cat}(z) \lor \text{Animal}(z) \\
& \neg \text{Kills}(\text{Curiosity}, \text{Tuna})
\end{align*}
\]

(D is a placeholder for the dog’s unknown name (i.e. Skolem symbol/function). Think of D like “JohnDoe”)

---

**Proof by Resolution**

---

**Completeness**

Resolution with unification applied to clausal form, is a complete inference procedure.

In practice, still significant search problem!

Many different search strategies: resolution strategies.

---

**Strategies for Selecting Clauses**

**unit-preference strategy:** Give preference to resolutions involving the clauses with the smallest number of literals.

**set-of-support strategy:** Try to resolve with the negated theorem or a clause generated by resolution from that clause.

**subsumption:** Eliminates all sentences that are subsumed (i.e., more specific than) an existing sentence in the KB.

May still require exponential time.

---

**Proofs can be Lengthy**

A relatively straightforward KB can quickly overwhelm general resolution methods.

Resolution strategies reduce the problem somewhat, but not completely.

As a consequence, many practical Knowledge Representation formalisms in AI use a restricted form and specialized inference.

- Logic programming (Prolog)
- Production systems
- Frame systems and semantic networks
- Description logics

---

**Successes in Rule-Based Reasoning**

**Expert systems**

- **DENDRAL** (Buchanan et al., 1969)
- **MYCIN** (Feigenbaum, Buchanan, Shortliffe)
- **PROSPECTOR** (Duda et al., 1979)
- **R1** (McDermott, 1982)
Successes in Rule-Based Reasoning

- **DENDRAL** (Buchanan et al., 1969)
  - Infers molecular structure from the information provided by a mass spectrometer
  - Generate-and-test method
  - If there are peaks at \( x_1 \) and \( x_2 \) s.t.
    \[
    x_1 + x_2 = M + 28
    \]
    \[
    x_1 - 28 \text{ is a high peak}
    \]
    \[
    x_2 - 28 \text{ is a high peak}
    \]
  - At least one of \( x_1 \) and \( x_2 \) is high
  - then there is a ketone subgroup

- **MYCIN** (Feigenbaum, Buchanan, Shortliffe)
  - Diagnosis of blood infections
  - 450 rules; performs as well as experts
  - Incorporated certainty factors
  - If: (1) the stain of the organism is gram positive, and
  - (2) the morphology of the organism is coccus, and
  - (3) the growth conformation of the organism is clumps,
    then there is suggestive evidence (0.7) that the identity of the organism is staphylococcus.

- **PROSPECTOR** (Duda et al., 1979)
  - Correctly recommended exploratory drilling at geological site
  - Rule-based system founded on probability theory

- **R1** (McDermott, 1982)
  - Designs configurations of computer components
  - About 10,000 rules
  - Uses meta-rules to change context
  - If: current context is ?x
  - then: deactivate ?x context
  - and activate ?y context

Cognitive Modeling with Rule-Based Systems

SOAR is a general architecture for building intelligent systems.
- Long term memory consists of rules
- Working memory describes current state
- All problem solving, including deciding what rule to execute, is state space search
- Successful rule sequences are chunked into new rules
- Control strategy embodied in terms of meta-rules

Properties of Knowledge-Based Systems

**Advantages**
1. Expressibility*
2. Simplicity of inference procedures*
3. Modifiability*
4. Explainability
5. Machine readability
6. Parallelism*

**Disadvantages**
1. Difficulties in expressibility
2. Undesirable interactions among rules
3. Non-transparent behavior
4. Difficult debugging
5. Slow
6. Where does the knowledge base come from???