Review 2: Logic-Based Inference

- **Deductive Inference in FOPL** (Chapter 9, Sections 10.3, 10.4)
  - Convert first order sentences to clause form
    - Definition of clauses
    - Conversion procedure
      step 1: Eliminate implication and equivalence symbols
      step 2: Move all negation symbols to individual predicates
      step 3: Eliminate all existential quantifiers (Skolemization)
      step 4: Eliminate all universally quantifiers
      step 5: Convert the sentence to conjunctive normal form
      step 6: use parenthesis to separate all disjunctions, then drop all v's and ^'s
  - Unification (obtain mgu θ)
    - Two terms x and y can be unified only if one of them, say x, is a variable and x does not appear anywhere in y. Then x/y is added into the substitution θ.
    - When one binding variable/term is found, apply it to the remainders of both argument lists and to previously found bindings before proceeding to unify other arguments
    - Two argument lists of the same predicate are unifiable if every corresponding pair of terms, one from each list, is unifiable
  - Resolution .
    - Two clauses C1 and C2 can be resolved if one contain literal P and the other contains ~P and the argument lists of P and ~P can be unified with mgu θ
    - The resulting clause (resolvant) is composed of all literals of C1 and C2 except P and ~P, subject to variable substitution according to θ.
  - Resolution Refutation
    - Write the axioms as FOL sentences and convert them into clause form
    - Write the goal (theorem) as a FOL sentence
    - Negate the goal and convert it to clause form
    - Select a pair of clauses for resolution which are
      i) resolvable, and ii) promising toward deriving a null clause,
    - Inference stops when a null clause is derived
  - Control strategies
    - Depth-first: incomplete
    - Breadth-first: complete
    - Unit resolution (one of the two parent clauses is a singleton clause): incomplete in general but complete for Horn clauses
    - Input resolution (one of the two parent clauses is from the original set of clauses): incomplete in general but complete for Horn clauses
    - Set of support: complete
    - Linear resolution (one of the two parent clauses is either from the original set of clauses or is an ancestor of the other parent clause): complete
  - Horn clauses and logic programming
    - Definition of Horn clauses
    - Advantages and limitations of using Horn clauses
    - Logic programming as a general purpose programming language (viewing and resolution as function calls answer extraction)
    - Features of Prolog (Horn clauses, top-down/left-right, depth-first plus backtracking)

- Advantages
  - Clearly defined semantics (least ambiguous
  - Expressiveness
  - Natural for some domains
- Disadvantages
  - Semantics is too rigid
  - Inefficiency (inference is NP}-hard with complete control strategies; semi-decidable)
  - Unnatural for many domains

- **Production (Rule-Based) Systems**] (Section 10.5)
  - System components: WM, rule base, inference engine (rule interpreter)
  - Inference procedure
    - Cycle of three phases: match, conflict-resolution, act/fire
    - Forward and backward inference
  - Conflict resolution
    - conflict set
    - conflict resolution policies (refraction, specificity, recency, priority/rule-ordering)
  - Advantages
    - Simplicity (for both language and inference)
    - Efficiency
    - Modularity (easy for KB maintenance)
    - Natural for many application domains
  - Disadvantages
    - No clearly defined semantics (based on informal understanding)
    - Incomplete inference procedure
    - Unpredictable side effects of ordering of rule applications
    - Less expressive (may not be suitable for some applications)

- **Structured representation**
  - Semantics (associative) networks
    - labeled nodes: objects, classes, concepts
    - Labeled directed links: relations (associations) between nodes
    - Reasoning about associations (marker passing and spreading activation)
  - ISA hierarchy and property inheritance
    - Super/subclass and instance/class relation
    - Inference by inheritance
    - Multiple inheritance (from different parents, from ancestors of different distances)
    - Exceptions in inheritance
  - Frame Systems
    - Definition (stereotypical views of the world; record like structure)
    - Slots, their values and facets
    - Procedural attachment and how they work (if-added, if-needed, if-updated)
    - Frames from different perspectives

- **Default reasoning**
  - Definition (inference is drawn in the absence of info to the contrary) and examples

- Default reasoning is non-monotonic, and it totally undecidable
- How production systems and semantic networks (and frame systems) handle simple default reasoning

Review 3: Abduction, Uncertainty, and Probabilistic Reasoning

- **Abduction**
  - Definition
  - Difference between abduction, deduction, and induction
  - Characteristics of abductive inference
    - Inference results are hypotheses, not theorems (may be false)
    - There may be multiple plausible hypotheses
    - Reasoning is often a hypothesize-and-test cycle
    - Reasoning is non-monotonic
    - Sources of uncertainty (uncertain data, knowledge, and inference)

- **Simple Bayesian approach to evidential/diagnostic reasoning**
  - Bayes' theorem
  - Conditional independence (and evidence) and single fault assumptions
  - Methods for computing posterior probability and relative likelihood of a hypothesis, given some evidence

$$P(H_i \mid E_1,\ldots,E_l) = \frac{P(E_1,\ldots E_l \mid H_i)P(H_i)}{P(E_1,\ldots E_l)}$$

$$rel(H_i \mid E_1,\ldots,E_l) = P(E_1,\ldots,E_l \mid H_i)P(H_i) = P(H_i)\prod_{j=1}^{l} P(E_j \mid H_i)$$

$$P(H_i \mid E_1,\ldots,E_l) = \frac{rel(H_i \mid E_1,\ldots,E_l)}{\sum_{k=1}^{n} rel(H_k \mid E_1,\ldots,E_l)} = \frac{P(H_i)\prod_{j=1}^{l} P(E_j \mid H_i)}{\sum_{k=1}^{n} P(H_k)\prod_{j=1}^{l} P(E_j \mid H_k)}$$

  - Evidence accumulation

$$rel(H_i \mid E_1,\ldots,E_l,\sim E_{l+1}) = (1 - P(E_{l+1} \mid H_i))rel(H_i \mid E_1,\ldots,E_l)$$

$$rel(H_i \mid E_1,\ldots,E_l,E_{l+1}) = P(E_{l+1} \mid H_i)rel(H_i \mid E_1,\ldots,E_l)$$

  - Limitation
    - Assumptions unreasonable for many problems
    - Not suitable for multi-fault problems
    - Can not represent causal chaining

- **Bayesian belief networks (BBN)**
  - Integration of probability theory and causal networks
  - Definition of BBN (DAG and CPD).

    $P(x_i \mid p_i)$ where p_i is the set of all parent nodes of x_i
  - Independence assumption

$$P(x_i \mid p_i, q) = P(x_i \mid p_i)$$

  - Computing joint probability distribution

$$P(x_1,\ldots,x_n) = \prod_{i=1}^{n} P(x_i \mid p_i)$$

- Inference (e.g., belief update, MAP) is NP}-complete (exponential time)
- BBN of noise-or gate (advantages and limitations)
- Learning BBN from case data (difficulty in learning the DAG)

- **Dempster-Shafer theory (for representing ignorance)**
  - Difference between probability of an event and ignorance
  - How to represent uncommitted belief (ignorance)
    - Lattice of subsets of frame of discernment
    - Basic probability assignment (function m(S))
    - Bel(S), Pls(S), and belief interval
  - Problem with this theory (high complexity)

- **Fuzzy set theory (for representing vague linguistic terms)**
  - Difference between fuzzy sets and ordinary sets
  - Fuzzy membership functions
  - Rules for fuzzy logic connectives
  - Problems with fuzzy logic (comparing with probability theory)

- **Uncertainty in rule-based system (certainty factors in MYCIN)**
  - CF of WM elements
  - CF of rules
  - CF propagation
  - Problems with CF