Propositional Logic

Operators: $\neg A$ (negation), $A \wedge B$ (conjunction/and), $A \vee B$ (disjunction/or), $A \to B$ (implication; $\neg A \vee B$), $A \leftrightarrow B$ (bi-implication).

τ **satisfies** A iff $\bar{\tau}(A)$ is true. τ satisfies Φ iff τ satisfies all formulas in Φ.

A is a **logical consequence** of Φ ($\Phi \models A$) iff for all τ, if τ satisfies Φ, then it satisfies A.

Limited by boolean variables (cannot cross reference between individuals in a statement) and the lack of quantifiers (having to list all members to specify a property).

First-Order Logic

Defined by: Variables V, functions F, predicates P. **Terms**: variables or functions of terms. 0-ary functions are constant terms (cannot be quantified). **Vocabulary**: $\mathcal{L}$, a set of function and predicate symbols. **Formula**: *atomic* ($P(t_1, \dots, t_n)$ where t_i are terms), or formulas combined with propositional operators, or $\exists$ and $\forall$ quantifiers.

Converting from English: things become constants, types/properties become unary predicates, relationships become binary (or more) predicates, associations become functions.

Structure: an $\mathcal{L}$-structure $\mathcal{M}$ contains the **universe** $M \neq \varnothing$, function extensions $f^{\mathcal{M}}\colon M^n \mapsto M$ for each $f \in \mathcal{L}$ (specified as individual mappings), predicate extensions $P^{\mathcal{M}} \subseteq M^n$ for each $P \in \mathcal{L}$ (specified as sets of n-tuples $\langle A, B, \dots \rangle$ for which the predicate is true). **Object assignment**: an object assignment σ for $\mathcal{M}$ is a mapping from a set of variables to the universe of M. Note $\sigma(m/x)$ is an assignment mapping x to $m \in M$ (for quantifiers).

Satisfaction: $\mathcal{M}$ is a model of C under σ (denoted $\mathcal{M} \models C[\sigma]$) if C is true, under the definitions and variable mappings of $\mathcal{M}$ and σ.

x is **bounded** in A if it only exists in A under a quantifier; otherwise it is **free**. If σ and σ' have the same assignment for all free variables of A, then $\mathcal{M} \models A[\sigma] \iff \mathcal{M} \models A[\sigma']$. A is a **sentence** if it is **closed** (no free variables). For sentences, σ is irrelevant and can be dropped. $\mathcal{M}$ is a **model** of Φ ($\mathcal{M} \models \Phi$) if it satisfies all sentences in Φ. Φ is **satisfiable** if there exists $\mathcal{M}$ that models Φ.

A is a **logical consequence** of Φ ($\Phi \models A$) iff for every $\mathcal{M}$, $\mathcal{M} \models \Phi \implies \mathcal{M} \models A$. If $\Phi \models A$, then $\nexists \mathcal{M}$ s.t. $\mathcal{M} \models \Phi \cup \{\neg A\}$.

Knowledge base: a collection of sentences representing the agent's beliefs, can be used for inference about implicit knowledge through proof procedures. Procedures are **sound** if it only produces logical consequences of the KB, and **complete** if it can produce all logical consequences of the KB.

Resolution by refutation: a sound and complete proof procedure; first assume $\neg A$ (refutation), then convert $\neg A$ and KB to a clausal theory C, and resolve clauses in C until the empty clause is reached, where we conclude A is true.

Clausal theory: a set (conjunction) of **clauses** that must all be true; each clause is a disjunction of **literals** (at least one is true); each literal is either an atomic formula or a negated atomic formula.

Resolution: $\dfrac{a_1 \vee \dots \vee a_n \vee c \qquad b_1 \vee \dots \vee b_m \vee \neg c}{a_1 \vee \dots \vee a_n \vee b_1 \vee \dots \vee b_m}$

Conversion to clausal form:

1. Convert implications: $A \to B$ to $\neg A \vee B$.
2. Move negations inward and simplify double negations: $\neg(A \wedge B) \iff \neg A \vee \neg B$, $\neg(A \vee B) \iff \neg A \wedge \neg B$, $\neg \forall x A \iff \exists x \neg A$, $\neg \exists x A \iff \forall x \neg A$.
3. Variable standardization: rename so that each quantified variable is unique.

4. Skolemization: remove $\exists$ by replacing the variable quantified with a new unique constant $\exists x P(x) \iff P(\boldsymbol{a})$, or unique function which mentions every variable that scopes the existential $\forall x \exists y P(x, y) \iff \forall x P(x, g(x))$.
5. Prenex form: take all quantifiers (only $\forall$ at this point) to the front: $\forall x P \wedge Q \iff \forall x(P \wedge Q)$, $\forall x P \vee Q \iff \forall x(P \vee Q)$.
6. Distribute $\vee$ over $\wedge$: $A \vee (B \wedge C) \iff (A \vee B) \wedge (A \vee C)$.
7. Convert to clauses: remove all $\forall$ quantifiers (implicit) and break apart $\wedge$.

Resolution is **refutation complete**: it can eventually prove that a clausal theory is unsatisfiable, but may not terminate when it is. In general first-order unsatisfiability is semi-decidable (there exists an algorithm that correctly gives positive answers), but not decidable.