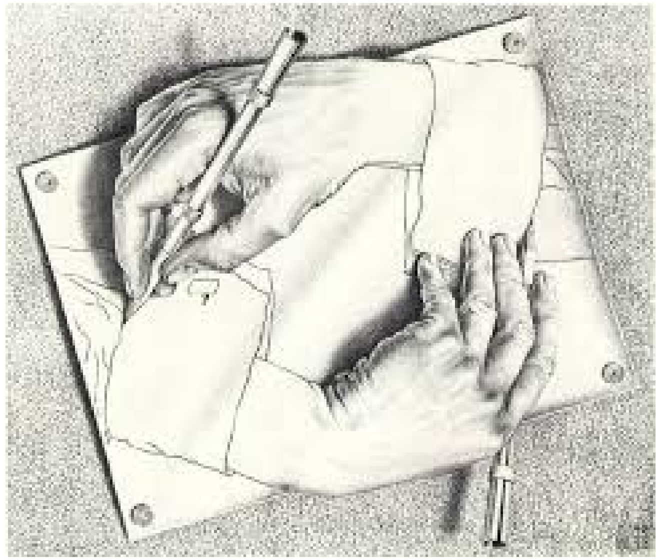
Logic: resolution



Resolution

Recall: First-order logic includes non-Horn clauses

$$\forall x\, \mathsf{Student}(x) \rightarrow \exists y\, \mathsf{Knows}(x, y)$$

High-level strategy (same as in propositional logic):

- Convert all formulas to CNF
- Repeatedly apply resolution rule

- To go beyond Horn clauses, we will develop a single resolution rule which is sound and complete.
- The high-level strategy is the same as propositional logic: convert to CNF and apply resolution.

Conversion to CNF

Input:

$$\forall x\,(\forall y\,\mathsf{Animal}(y) \to \mathsf{Loves}(x,y)) \to \exists y\,\mathsf{Loves}(y,x)$$

Output:

$$(\mathsf{Animal}(Y(x)) \vee \mathsf{Loves}(Z(x),x)) \wedge (\neg\mathsf{Loves}(x,Y(x)) \vee \mathsf{Loves}(Z(x),x))$$

New to first-order logic:

- All variables (e.g., x) have universal quantifiers by default
- Introduce **Skolem functions** (e.g., $Y(x)$) to represent existential quantified variables

- Consider the logical formula corresponding to *Everyone who loves all animals is loved by someone*. The slide shows the desired output, which looks like a CNF formula in propositional logic, but there are two differences: there are variables (e.g., x) and functions of variables (e.g., $Y(x)$). The variables are assumed to be universally quantified over, and the functions are called **Skolem functions** and stand for a property of the variable.

Conversion to CNF (part 1)

Anyone who likes all animals is liked by someone.

Input:

$\forall x\,(\forall y\,\text{Animal}(y) \rightarrow \text{Loves}(x, y)) \rightarrow \exists y\,\text{Loves}(y, x)$

Eliminate implications (old):

$\forall x\,\neg(\forall y\,\neg\text{Animal}(y) \vee \text{Loves}(x, y)) \vee \exists y\,\text{Loves}(y, x)$

Push $\neg$ inwards, eliminate double negation (old):

$\forall x\,(\exists y\,\text{Animal}(y) \wedge \neg\text{Loves}(x, y)) \vee \exists y\,\text{Loves}(y, x)$

Standardize variables (**new**):

$\forall x\,(\exists y\,\text{Animal}(y) \wedge \neg\text{Loves}(x, y)) \vee \exists z\,\text{Loves}(z, x)$

- We start by eliminating implications, pushing negation inside, and eliminating double negation, which is all old.
- The first thing new to first-order logic is standardization of variables. Note that in $\exists x\, P(x) \wedge \exists x\, Q(x)$, there are two instances of x whose scopes don't overlap. To make this clearer, we will convert this into $\exists x\, P(x) \wedge \exists y\, Q(y)$. This sets the stage for when we will drop the quantifiers on the variables.

Conversion to CNF (part 2)

$\forall x \, (\exists y \, \text{Animal}(y) \wedge \neg \text{Loves}(x, y)) \vee \exists z \, \text{Loves}(z, x)$

Replace existentially quantified variables with Skolem functions (**new**):

$\forall x \, [\text{Animal}(Y(x)) \wedge \neg \text{Loves}(x, Y(x))] \vee \text{Loves}(Z(x), x)$

Distribute $\vee$ over $\wedge$ (old):

$\forall x \, [\text{Animal}(Y(x)) \vee \text{Loves}(Z(x), x)] \wedge [\neg \text{Loves}(x, Y(x)) \vee \text{Loves}(Z(x), x)]$

Remove universal quantifiers (**new**):

$[\text{Animal}(Y(x)) \vee \text{Loves}(Z(x), x)] \wedge [\neg \text{Loves}(x, Y(x)) \vee \text{Loves}(Z(x), x)]$

- The next step is to remove existential variables by replacing them with Skolem functions. This is perhaps the most non-trivial part of the process. Consider the formula: $\forall x\,\exists y\,P(x,y)$. Here, y is existentially quantified and depends on x. So we can mark this dependence explicitly by setting $y = Y(x)$. Then the formula becomes $\forall x\,P(x, Y(x))$. You can even think of the function Y as being existentially quantified over outside the $\forall x$.
- Next, we distribute disjunction over conjunction as before.
- Finally, we simply drop all universal quantifiers. Because those are the only quantifiers left, there is no ambiguity.
- The final CNF formula can be difficult to interpret, but we can be assured that the final formula captures exactly the same information as the original formula.

Resolution

Definition: resolution rule (first-order logic)

$$\frac{f_1 \vee \cdots \vee f_n \vee p, \quad \neg q \vee g_1 \vee \cdots \vee g_m}{\text{Subst}[\theta, f_1 \vee \cdots \vee f_n \vee g_1 \vee \cdots \vee g_m]}$$

where $\theta = \text{Unify}[p, q]$.

Example: resolution

$$\frac{\text{Animal}(Y(x)) \vee \text{Loves}(Z(x), x), \quad \neg\text{Loves}(u, v) \vee \text{Feeds}(u, v)}{\text{Animal}(Y(x)) \vee \text{Feeds}(Z(x), x)}$$

Substitution: $\theta = \{u/Z(x), v/x\}$.

- After convering all formulas to CNF, then we can apply the resolution rule, which is generalized to first-order logic. This means that instead of doing exact matching of a literal p, we unify atomic formulas p and q, and then apply the resulting substitution θ on the conclusion.