These complexity results motivate the study of subclasses of formulas (*fragments*) of first-order logic

*Q*: Can you think of any fragments of first-order logic for which validity is decidable?

## Some Decidable Fragments

Some decidable fragments:

- *Monadic class*: no function symbols, all predicates unary; validity is NEXPTIME-complete.
- Variable-free formulas without equality: satisfiability is NP-complete. (why?)
- Variable-free Horn clauses (clauses with at most one positive atom): entailment is decidable in linear time.
- Finite model checking is decidable in time polynomial in the size of the structure and the formula.

## 3.5 Normal Forms and Skolemization (Traditional)

Study of normal forms motivated by

- reduction of logical concepts,
- efficient data structures for theorem proving.

The main problem in first-order logic is the treatment of quantifiers. The subsequent normal form transformations are intended to eliminate many of them.

### Prenex Normal Form

*Prenex formulas* have the form

$$Q_1x_1 \ldots Q_nx_n\ F,$$

where $F$ is quantifier-free and $Q_i \in \{\forall, \exists\}$; we call $Q_1x_1 \ldots Q_nx_n$ the *quantifier prefix* and $F$ the *matrix* of the formula.

Computing prenex normal form by the rewrite relation $\Rightarrow_P$:

$$\begin{array}{rcll}
(F \leftrightarrow G) & \Rightarrow_P & (F \rightarrow G) \wedge (G \rightarrow F) & \\
\neg QxF & \Rightarrow_P & \overline{Q}x\neg F & (\neg Q) \\
(QxF\ \rho\ G) & \Rightarrow_P & Qy(F[y/x]\ \rho\ G),\ y \text{ fresh},\ \rho \in \{\wedge, \vee\} & \\
(QxF \rightarrow G) & \Rightarrow_P & \overline{Q}y(F[y/x] \rightarrow G),\ y \text{ fresh} & \\
(F\ \rho\ QxG) & \Rightarrow_P & Qy(F\ \rho\ G[y/x]),\ y \text{ fresh},\ \rho \in \{\wedge, \vee, \rightarrow\} &
\end{array}$$

Here $\overline{Q}$ denotes the quantifier *dual* to $Q$, i.e., $\overline{\forall} = \exists$ and $\overline{\exists} = \forall$.

### Skolemization

**Intuition:** replacement of $\exists y$ by a concrete choice function computing $y$ from all the arguments $y$ depends on.

Transformation $\Rightarrow_S$ (to be applied outermost, *not* in subformulas):

$$\forall x_1, \ldots, x_n \exists y F \ \Rightarrow_S\ \forall x_1, \ldots, x_n F[f(x_1, \ldots, x_n)/y]$$

where $f$, where $\mathsf{arity}(f) = n$, is a new function symbol (*Skolem function*).

**Together:** $F \overset{*}{\Rightarrow}_P \underbrace{G}_{\text{prenex}} \overset{*}{\Rightarrow}_S \underbrace{H}_{\text{prenex, no } \exists}$

**Theorem 3.9** *Let $F$, $G$, and $H$ as defined above and closed. Then* *(i)* $F$ *and* $G$ *are equivalent.*

*(ii)* $H \models G$ *but the converse is not true in general.*

*(iii)* $G$ *satisfiable (w. r. t.* $\Sigma$*-Alg)* $\Leftrightarrow$ $H$ *satisfiable (w. r. t.* $\Sigma'$*-Alg) where* $\Sigma' = (\Omega \cup SKF, \Pi)$*, if* $\Sigma = (\Omega, \Pi)$*.*

### Clausal Normal Form (Conjunctive Normal Form)

$$
\begin{array}{rcl}
(F \leftrightarrow G) & \Rightarrow_K & (F \rightarrow G) \wedge (G \rightarrow F) \\
(F \rightarrow G) & \Rightarrow_K & (\neg F \vee G) \\
\neg(F \vee G) & \Rightarrow_K & (\neg F \wedge \neg G) \\
\neg(F \wedge G) & \Rightarrow_K & (\neg F \vee \neg G) \\
\neg\neg F & \Rightarrow_K & F \\
(F \wedge G) \vee H & \Rightarrow_K & (F \vee H) \wedge (G \vee H) \\
(F \wedge \top) & \Rightarrow_K & F \\
(F \wedge \bot) & \Rightarrow_K & \bot \\
(F \vee \top) & \Rightarrow_K & \top \\
(F \vee \bot) & \Rightarrow_K & F
\end{array}
$$

These rules are to be applied modulo associativity and commutativity of $\wedge$ and $\vee$. The first five rules, plus the rule $(\neg Q)$, compute the *negation normal form* (NNF) of a formula.

### The Complete Picture

$$
\begin{array}{lll}
F & \overset{*}{\Rightarrow}_P & Q_1 y_1 \ldots Q_n y_n \; G \qquad\qquad\qquad (G \text{ quantifier-free}) \\
 & \overset{*}{\Rightarrow}_S & \forall x_1, \ldots, x_m \; H \qquad\qquad (m \leq n,\ H \text{ quantifier-free}) \\
 & \overset{*}{\Rightarrow}_K & \underbrace{\underbrace{\forall x_1, \ldots, x_m}_{\text{leave out}} \; \bigwedge_{i=1}^{k} \underbrace{\bigvee_{j=1}^{n_i} L_{ij}}_{\text{clauses } C_i}}_{F'}
\end{array}
$$

$N = \{C_1, \ldots, C_k\}$ is called the *clausal (normal) form* (CNF) of $F$.
*Note:* the variables in the clauses are implicitly universally quantified.

**Theorem 3.10** *Let* $F$ *be closed. Then* $F' \models F$*. (The converse is not true in general.)*

**Theorem 3.11** *Let* $F$ *be closed. Then* $F$ *is satisfiable iff* $F'$ *is satisfiable iff* $N$ *is satisfiable*