

- fol with sorts
- fol with equality
- constraint satisfaction
- datalog
- description logics

- in commonsense reasoning, computer science and many applications sorts play an important role
- compare *every dog is an animal* with *every object in the domain that is a dog is also an animal*
- or the two formulas

$$\begin{aligned} & (\forall X : \textit{number})(\textit{odd}(X) \Rightarrow \textit{odd}(\textit{succ}(\textit{succ}(X)))) \\ & (\forall X)(\textit{number}(X) \wedge \textit{odd}(X) \Rightarrow \textit{odd}(\textit{succ}(\textit{succ}(X)))) \end{aligned}$$

- already in 1885 Pierce has suggested to annotate quantified variables with **sorts** designating sets of objects

- a **first order language with sorts** is a FOL language with alphabet \mathcal{A} together with a function

$$\text{sort} : \mathcal{A}_V \longrightarrow 2^{\mathcal{A}_S}$$

where \mathcal{A}_V is the set of variables in \mathcal{A} ,

and $\mathcal{A}_S \subseteq \mathcal{A}_R$ is a finite set of unary (monadic) relations in the alphabet

- each member of \mathcal{A}_S is called **base sort**
- a **sort** s is a non-empty set of base sorts, $s \in 2^{\mathcal{A}_S}$

- variables are **annotated** by their sort

$$X : s \text{ if } \text{sort}(X) = s$$

- finally we assume that for every sort s there are countably many variables $X : s$
- to assign meaning to sorted formulas we extend the notion of interpretation to sorts
- let D be the domain of an interpretation I , then I maps each sort $s = \{p_1, p_2, \dots, p_n\}$ to

$$s^I = p_1^I \cap p_2^I \dots \cap p_n^I \subseteq D$$

- a **sorted state** σ is a state in which all variables $X : s$ are such that $\sigma(X) \in s^I$
- the application of a sorted state to a term is the same as in traditional FOL

- let I be an interpretation, and σ a sorted state, then the meaning of a formula F in a sorted language is defined inductively as in FOL, with the difference that

$$I, \sigma \models (\exists X : s)F \quad \text{iff there exists } e \in s^I \text{ such that } I, \sigma\{X \setminus e\} \models F$$
$$I, \sigma \models (\forall X : s)F \quad \text{iff for all } e \in s^I \text{ it holds that } I, \sigma\{X \setminus e\} \models F$$

- the model and entailment concepts are also defined as in FOL

- FOL with sorts seems to be a **generalization** of standard FOL
- but it is possible to **transform** each valid formula in a sorted language in an equivalent unsorted FOL formula

$$t(p(t_1, \dots, t_n)) := p(t_1, \dots, t_n)$$

$$t(F \wedge G) := t(F) \wedge t(G)$$

⋮

$$t((\forall X : s)F) := (\forall X)(p_1(X) \wedge p_2(X) \wedge \dots \wedge p_n(X) \Rightarrow t(F))$$

if $s = \{p_1, p_2, \dots, p_n\}$

$$t((\exists X : s)F) := (\exists X)(p_1(X) \wedge p_2(X) \wedge \dots \wedge p_n(X) \Rightarrow t(F))$$

if $s = \{p_1, p_2, \dots, p_n\}$

- transformation $t(\cdot)$ can be computed **linearly**, and also with a linear expansion of the formula
- it can be shown that every model of the sorted formula is a model of the transformed formula
- thus the **expressive power** of sorted and unsorted FOL is identical
- however, in a calculus where the sort information has been built into the deductive machinery, computations **may be considerable faster** [Wei96]

- the equality relation plays an important role in mathematics, artificial intelligence, operations research, etc
- we now describe how to introduce it in the framework of FOL
- we extend the alphabet \mathcal{A} with a binary relation symbol \approx typically written infix
- an **equation** is an expression of the form $s \approx t$, with s, t terms
- an **equational system** is a set of universally closed equations

- example: specification of a group \mathcal{E}_G

$$(\forall X, Y, Z)(X \cdot Y) \cdot Z \approx X \cdot (Y \cdot Z)$$

$$(\forall X)1 \cdot X \approx X$$

$$(\forall X)X \cdot 1 \approx X$$

$$(\forall X)X^{-1} \cdot X \approx 1$$

$$(\forall X)X \cdot X^{-1} \approx 1$$

- so far, the equality relation is just any ordinary relation symbol
- but we usually expect equality to have **properties** like symmetry, transitivity, reflexivity and substitutivity

$$\begin{aligned} & (\forall X) X \approx X \\ & (\forall X, Y) X \approx Y \Rightarrow Y \approx X \\ & (\forall X, Y, Z) X \approx Y \wedge Y \approx Z \Rightarrow X \approx Z \\ & (\forall X_1, \dots, X_n, Y_1, \dots, Y_n) (\bigwedge_{i=1}^n X_i \approx Y_i \Rightarrow f(X_1, \dots, X_n) \approx f(Y_1, \dots, Y_n)) \\ & (\forall X_1, \dots, X_n, Y_1, \dots, Y_n) (\bigwedge_{i=1}^n X_i \approx Y_i \wedge p(X_1, \dots, X_n) \Rightarrow p(Y_1, \dots, Y_n)) \end{aligned}$$

- these properties are called the **axioms for equality**, and designated by \mathcal{E}_{\approx}
- as usual we are interested in the logical consequence of an equational systems \mathcal{E} , ie the formulas F such that

$$\mathcal{E} \cup \mathcal{E}_{\approx} \models F$$

- example: every group satisfying $(\forall X)X \cdot X \approx 1$ is commutative

$$\mathcal{E}_G \cup \mathcal{E}_{\approx} \cup \{(\forall X)X \cdot X \approx 1\} \models (\forall X, Y)X \cdot Y \approx Y \cdot X$$

- all these formulas are first order and in clause form, so we can apply resolution
- however, it is obvious that the **search space** generated by all possible applications of resolution to the theory is extremely large
- also, it does not correspond to the way mathematicians would solve the problem

- so resolution is clearly **impractical** for equality
- a solution is to **extend the calculus** by adding a new inference rule well-suited for equality, or to **modify resolution itself** to include the axioms of equality in the unification step
- **paramodulation** is one of the solutions in the former sense
- let $L[s]$ be a literal that contains an occurrence of s , and $L[s \setminus t]$ the same literal where this occurrence has been replaced by t
- if $C_1 = \{L_1[s], L_2, \dots, L_n\}$ and $C_2 = \{l \approx r, K_1, \dots, K_m\}$ are two clauses standardized apart and if s and l are unifiable with mgu θ , then

$$\{L_1[s \setminus r]\theta, L_2\theta, \dots, L_n\theta, K_1\theta, \dots, K_m\theta\}$$

is the **paramodulation** of C_1 and C_2

- **Theorem 29.** *if a set of clauses $\mathcal{E} \cup \mathcal{E}_{\approx} \cup \{\neg s \approx t\}$ is unsatisfiable, then there is a refutation of $\mathcal{E} \cup \{X \approx X, \neg s \approx t\}$ using paramodulation, resolution, and factoring.*
- this means that every axiom of equality, except of reflexivity, can be built into paramodulation
- **Exercise:** show that the axioms of symmetry and transitivity can be derived from the axioms of reflexivity and f-substitutivity
- **Exercise:** show that each resolution step using either the axiom of symmetry, transitivity, or substitutivity can be simulated by resolution and paramodulation. In the simulation, resolution cannot be used with the mentioned equational axioms.

- example: in order to show that

$$\mathcal{E}_G \cup \mathcal{E}_{\approx} \cup \{(\forall X)X \cdot X \approx 1\} \models (\forall X, Y)X \cdot Y \approx Y \cdot X$$

from theorem 3 we can show that

$$\bigwedge (\mathcal{E} \cup \mathcal{E}_{\approx} \cup \{(\forall X)X \cdot X \approx 1\}) \Rightarrow (\forall X, Y)X \cdot Y \approx Y \cdot X$$

is valid,

and thanks to theorem 4 this is equivalent to

$$\neg(\bigwedge (\mathcal{E} \cup \mathcal{E}_{\approx} \cup \{(\forall X)X \cdot X \approx 1\}) \Rightarrow (\forall X, Y)X \cdot Y \approx Y \cdot X))$$

being unsatisfiable

So when this last formula is translated into clausal form we get

$$\mathcal{E} \cup \mathcal{E}_{\approx} \cup \{X \cdot X \approx 1\} \cup \{a \cdot b \neq b \cdot a\}$$

- example (cont.): we can now try to refute this set of formulas

- | | | |
|----|---|---|
| 1) | $a \cdot b \not\approx \underline{b} \cdot a$ | initial query |
| 2) | $1 \cdot X_1 \approx X_1$ | left unit |
| 3) | $\underline{X_2} \approx X_2$ | reflexivity |
| 4) | $\underline{X_1} \approx 1 \cdot X_1$ | paramodulation on 3) and 2) with $\theta = \{X_2 \setminus 1 \cdot X_1\}$ |
| 5) | $a \cdot b \not\approx (1 \cdot b) \cdot a$ | paramodulation on 1) and 4) with $\theta = \{X_1 \setminus b\}$ |
| 6) | $X_3 \cdot X_3 \approx 1$ | hypothesis |
| 7) | $\underline{X_4} \approx X_4$ | reflexivity |
| 8) | $1 \approx X_3 \cdot X_3$ | paramodulation on 7) and 6) with $\theta = \{X_4 \setminus X_3 \cdot X_3\}$ |
| 9) | $a \cdot b \not\approx ((X_3 \cdot X_3) \cdot b) \cdot \underline{a}$ | paramodulation on 5) and 8) with $\theta = \{\}$ |
| | ⋮ | right unit |
| k) | $a \cdot b \not\approx ((X_3 \cdot X_3) \cdot b) \cdot (a \cdot \underline{1})$ | paramodulation on 9) and ... |

• example (cont.):

$k)$	$a \cdot b \not\approx ((X_3 \cdot X_3) \cdot b) \cdot (a \cdot \underline{1})$	paramodulation on 9) and \dots
	\vdots	hypothesis
$k')$	$a \cdot b \not\approx \underline{((X_3 \cdot X_3) \cdot b) \cdot (a \cdot (X_4 \cdot X_4))}$	paramodulation on $k)$ and \dots
	\vdots	associativity
$k'')$	$a \cdot b \not\approx (X_3 \cdot \underline{((X_3 \cdot b) \cdot (a \cdot X_4))}) \cdot X_4$	paramodulation on $k')$ and \dots
	\vdots	hypothesis
$k''')$	$a \cdot b \not\approx \underline{(a \cdot 1)} \cdot b$	paramodulation on k'') and \dots
	\vdots	with $\theta = \{X_3 \setminus a, X_4 \setminus b\}$
$n)$	$a \cdot b \not\approx a \cdot b$	right unit
$n')$	$X_5 \approx X_5$	paramodulation on $k''')$ and \dots with $\theta = \{\}$
n'')	$\{\}$	reflexivity
		resolvent of $n)$ and $n')$

- the refutation is clumsy but it follows closely the way mathematicians transform equations under equality
- the search space is still huge, with many steps redundant and useless
- **Exercise:** Consider the set of clauses

$$\mathcal{F} = \{\{p(f(a)), q(y), r(b)\}, \{\neg p(f(b))\}, \{\neg q(a)\}, \{\neg r(a)\}\}$$

and the equational system $\mathcal{E} = \{f(X) \approx X, a \approx b\}$. Show by resolution and paramodulation that $\mathcal{F} \cup \mathcal{E}$ is unsatisfiable.

- **unification theory** is another calculus based on resolution for equality theories,
- in which equality is taking care within the unification step
- the general form of unification problems is the question whether

$$\mathcal{E} \cup \mathcal{E}_{\approx} \models (\exists X_1, \dots, X_n) t \approx s$$

- for example,

$$\mathcal{E} \cup \mathcal{E}_{\approx} \models (\exists X, Y) f(X, g(a, b)) \approx f(g(Y, b), X)$$

- if \mathcal{E} is empty, then this decision problem is decidable in general
- for the example, the mgu $\theta = \{X \setminus g(a, b), Y \setminus a\}$ is the unique minimal solution that can be computed using several unification algorithms available

- if \mathcal{E} states that f is commutative, ie $\mathcal{E} = \{(\forall X, Y) f(X, Y) \approx f(Y, X)\}$ then θ is not minimal and mgu, since $\sigma = \{Y \setminus a\}$ is also a solution and more general
- in general, the existence of **at most one** mgu does not hold any more for unification under commutativity, but the number of solutions is **finite**

- if \mathcal{E} states that f is associative, ie $\mathcal{E} = \{(\forall X, Y, Z) f(X, f(Y, Z)) \approx f(f(X, Y), Z)\}$
- then θ is still a solution but also

$$\rho = \{X \setminus f(g(a, b), g(a, b)), Y \setminus a\}$$

- and neither θ is more general than ρ , nor ρ is more general than θ
- it is easy to prove that in general there can be **infinitely many** solutions in this case
- if \mathcal{E} states that f is both associative and commutative, then the situation changes again and the number of independent solutions is again **finite**
- in **unification theory** several equational systems are studied in order to know how many solutions can exist, if the decision problem is decidable, and in such cases which is the complexity of this problem.