Question 1. (3 points)

Determine the least general generalization of the following two assertions

1. *Superman is mortal or he is not a human.*
2. *Every human who smokes is mortal.*

by representing them as first-order logic clauses and computing their least general generalization with respect to the θ-subsumption order, and express the result in natural language.

Answer:

The first clause is

$$\text{mortal}(\mathsf{superman}) \vee \neg\text{human}(\mathsf{superman})$$

The second clause is

$$\text{human}(x) \wedge \text{smokes}(x) \rightarrow \text{mortal}(x)$$

which can be written as

$$\neg\text{human}(x) \vee \neg\text{smokes}(x) \vee \text{mortal}(x)$$

The LGG is

$$\neg\text{human}(\mathsf{lgg}(\mathsf{superman}, x)) \vee \text{mortal}(\mathsf{lgg}(\mathsf{superman}, x)) \quad (1)$$
$$\text{i.e.,} \quad (2)$$
$$\neg\text{human}(v_1) \vee \text{mortal}(v_1) \quad (3)$$
$$\text{i.e.,} \quad (4)$$
$$\text{human}(v_1) \rightarrow \text{mortal}(v_1) \quad (5)$$

i.e., *every human is mortal*.

Question 2. (2 points)

Let h, h' be FOL clauses and B a ground FOL conjunction. Show that if $h \subseteq_\theta h'$ then $h \subseteq_\theta^B h'$.

Answer:

$h \subseteq_\theta h'$ means that

$$\exists\theta : \mathsf{Lits}(h\theta) \subseteq \mathsf{Lits}(h') \quad (6)$$

$h \subseteq_\theta^B h'$ is defined as $h \subseteq_\theta B \rightarrow h'$, i.e.,

$$\exists\theta : \mathsf{Lits}(h\theta) \subseteq \mathsf{Lits}(B \rightarrow h')$$
$$\exists\theta : \mathsf{Lits}(h\theta) \subseteq \mathsf{Lits}(\neg B \vee h')$$
$$\exists\theta : \mathsf{Lits}(h\theta) \subseteq \mathsf{Lits}(\neg B) \cup \mathsf{Lits}(h') \quad (7)$$

where (7) is clearly implied by (6).

Question 3. (5 points)

Show that

$$h = \text{parent}(v_2, v_1) \wedge \text{male}(v_1) \rightarrow \text{son}(v_1, v_2)$$

and

$$g = \text{son}(v_1, v_2) \lor \neg\text{female}(\mathsf{a}) \lor \neg\text{parent}(\mathsf{a}, \mathsf{b}) \lor \neg\text{parent}(v_2, v_1) \lor \neg\text{male}(\mathsf{b}) \lor \neg\text{male}(v_1) \lor \neg\text{parent}(v_3, v_4) \lor \neg\text{parent}(\mathsf{b}, \mathsf{c}) \lor \neg\text{male}(v_4) \lor \neg\text{male}(\mathsf{c})$$

are equivalent relative to

$$B = \text{female}(\mathsf{a}) \land \text{parent}(\mathsf{a}, \mathsf{b}) \land \text{male}(\mathsf{b}) \land \text{parent}(\mathsf{b}, \mathsf{c}) \land \text{male}(\mathsf{c})$$

Answer:

$h \approx_\theta^B g$ iff $h \subseteq_\theta^B g$ and $g \subseteq_\theta^B h$.

$h \approx_\theta^B g$ because $h \subseteq_\theta g$, which is because $\mathsf{Lits}(h\theta) \subseteq \mathsf{Lits}(g)$ with $\theta = \emptyset$.

$g \approx_\theta^B h$ because $g \subseteq_\theta B \to h$, which is because $\mathsf{Lits}((B \to g)\theta) \subseteq \mathsf{Lits}(h)$ with $\theta = \{ v_3 \mapsto v_2, v_4 \mapsto v_1 \}$.

Question 4. (10 points)

Let

$$B = \text{half}(4, 2) \land \text{half}(2, 1) \land \text{int}(2) \land \text{int}(1)$$
$$x_1 = \text{even}(4)$$
$$x_2 = \text{even}(2)$$

1. Compute a least general generalization of x_1, x_2 observations relative to B .
2. Determine the reduction of the resulting clause relative to B and justify why it is indeed a reduction of it relative to B.

Answer:

1. $\text{rlgg}_B(x_1, x_2) = \mathsf{lgg}(B \to x_1, B \to x_2)$ where

$$B \to x_1 = \text{even}(4) \lor \neg\text{half}(4, 2) \lor \neg\text{half}(2, 1) \lor \neg\text{int}(2) \lor \neg\text{int}(1)$$
$$B \to x_2 = \text{even}(2) \lor \neg\text{half}(4, 2) \lor \neg\text{half}(2, 1) \lor \neg\text{int}(2) \lor \neg\text{int}(1)$$

	$\text{even}(4)$	$\neg\text{half}(4,2)$	$\neg\text{half}(2,1)$	$\neg\text{int}(2)$	$\neg\text{int}(1)$
$\text{even}(2)$	$\text{even}(v_1)$				
$\neg\text{half}(4,2)$		$\neg\text{half}(4,2)$	$\neg\text{half}(v_3,v_4)$		
$\neg\text{half}(2,1)$		$\neg\text{half}(v_1,v_2)$	$\neg\text{half}(2,1)$		
$\neg\text{int}(2)$				$\neg\text{int}(2)$	$\neg\text{int}(v_4)$
$\neg\text{int}(1)$				$\neg\text{int}(v_2)$	$\neg\text{int}(1)$

θ	σ	new variable
2	4	v_1
1	2	v_2
4	2	v_3
2	1	v_4

$$\mathsf{lgg} = \text{even}(v_1) \lor \neg\text{half}(4, 2) \lor \neg\text{half}(v_3, v_4) \lor \neg\text{half}(v_1, v_2) \lor \neg\text{half}(2, 1) \lor \neg\text{int}(2) \lor \neg\text{int}(v_4) \lor \neg\text{int}(v_2) \lor \neg\text{int}(1)$$

or

$$\text{even}(v_1) \leftarrow \text{half}(4, 2) \land \text{half}(v_3, v_4) \land \text{half}(v_1, v_2) \land \text{half}(2, 1) \land \text{int}(2) \land \text{int}(v_4) \land \text{int}(v_2) \land \text{int}(1)$$

2. Remove ground literals present in B obtaining a clause which is subsume-equivalent (relative to B) to the lgg.

$$\text{even}(v_1) \leftarrow \text{half}(v_3, v_4) \land \text{half}(v_1, v_2) \land \text{int}(v_4) \land \text{int}(v_2)$$

This is equivalent to

$$\text{even}(v_1) \leftarrow \text{half}(v_1, v_2) \land \text{int}(v_2)$$

The latter subsumes the former evidently as its is a subset of the former's literals. The former subsumes the latter under substitution $\theta = \{ v_3 \mapsto v_1, v_4 \mapsto v_2 \}$. The last clause is a reduction of the LGG as it is subsume-equivalent to it relative to B as shown above, and it cannot be reduced further.

Question 5. (10 points)

Let X contain Herbrand interpretations for a finite set of $\mathcal{P}$ predicates and a finite set $\mathcal{F}$ of functions, and the observation complexity n_X be the tuple $(|\mathcal{P}|, |\mathcal{F}|)$. Show that the hypothesis class st-CNF (i.e., conjunctions of FOL clauses with at most s literals and at most t term occurrences in each literal) is learnable online from X.

Answer:

We reduce online st-CNF learning from X to learning a propositional monotone conjunction from truth assignments. More precisely, let $c_1, c_2, \ldots, c_{n'}$ be all st clauses made using $\mathcal{P}$ and $\mathcal{F}$. We learn a monotone propositional conjunction on n' variables from $\{0,1\}^{n'}$; for each observation $x \in X$ we present the learner examples $x'(x) \in \{0,1\}^{n'}$ such that

$$x'^i(x) = 1 \text{ iff } x \models c_i$$

We know that the monotone propositional conjunction can be learned online from $\{0,1\}^{n'}$ with mistake bound $poly(n')$. To show online learnability st-CNF from X, we only need to show $n' \leq poly(n_X) = poly(|\mathcal{P}|, |\mathcal{F}|)$.

An st-clause has no more that st different variables. So the maximum number of different terms in an st-clause is $|\mathcal{F}| + st$. An atom consists of a single predicate ($|\mathcal{P}|$ different choices) and at most t terms (each from from $|\mathcal{F}| + st$ choices). So there are $|\mathcal{P}|(|\mathcal{F}| + st)^t$ different atoms, i.e. $2|\mathcal{P}|(|\mathcal{F}| + st)^t$ different literals.

An st-clause combines at most s literals so there are at most

$$\mathcal{O}\binom{2|\mathcal{P}|(|\mathcal{F}| + st)^t}{s} = poly(|\mathcal{P}|, |\mathcal{F}|)$$

different st-clauses.