Knowledge Representation and Reasoning

Exercise 2

Albert-Ludwigs-Universität Freiburg

Bernhard Nebel, Gregor Behnke, Thorsten Engesser, Rolf-David Bergdoll, Leonardo Mieschendahl, Johannes Herrmann

November 5th 2021

Exercise 2.1 (a)

Model of satisfiable Horn Formulae

Consider a satisfiable Horn formula ψ. Consider the interpretation in which a variable is true if and only if it is true in all models of ψ. Prove that this interpretation is also a model of ψ.

Idea: We examine each possible type of clause of a Horn formula and show that it is satisfied by this interpretation.

Exercise 2.1 (a)

Model of satisfiable Horn Formulae – Formalization

First we formalize the interpretation described in the exercise:
Let ψ be a satisfiable Horn formula and $\mathcal{I}$ be the following interpretation of the variables $\{X_1, \ldots, X_n\}$ occurring in ψ:

$$\mathcal{I}(X_i) = \begin{cases} \mathbf{T} & \text{if } \mathcal{I}'(X_i) = \mathbf{T} \text{ for all models } \mathcal{I}' \text{ of } \psi \\ \mathbf{F} & \text{otherwise} \end{cases}$$

Each clause of ψ contains either no positive literal (**goal clause**) or exactly one positive literal (**definite clause**). We show that $\mathcal{I}$ makes all clauses of ψ true.

Exercise 2.1 (a)

Models for satisfiable Horn Formulae – Goal Clauses

Let $\chi = \{\neg Y_1, \ldots, \neg Y_k\}$ be an arbitrary goal clause with $\{Y_1, \ldots, Y_k\} \subseteq \{X_1, \ldots, X_n\}$.

- Let us assume that χ is not satisfied under $\mathcal{I}$. For this to be the case, all Y_i have to be true under $\mathcal{I}$.
- Since $\mathcal{I}$ only assigns true, if the respective variable is true in all models of ψ, we get $\mathcal{I}'(Y_i) = \mathbf{T}$ for all models $\mathcal{I}'$ and all variables $Y_i \in \{Y_1, \ldots, Y_k\}$.
- As this makes χ false, there cannot be a model at all and ψ is unsatisfiable $\leadsto$ contradiction. Therefore, there has to be at least one Y_i with $\mathcal{I}(Y_i) = \mathbf{F}$. Consequently, χ is true under $\mathcal{I}$.

Exercise 2.1 (a)

Model for satisfiable Horn Formulae – Definite Clauses

Let $\chi = \{Y_1, \neg Y_2, \ldots, \neg Y_k\}$ be an arbitrary definite clause with $\{Y_1, \ldots, Y_k\} \subseteq \{X_1, \ldots, X_n\}$. Y_1 is the solitary positive literal (which might also be the only literal).

- If there is a Y_i with $i \geq 2$ and $\mathcal{I}(Y_i) = \mathbf{F}$ then χ is satisfied. If there is no such Y_i, again we have $\mathcal{I}'(Y_j) = \mathbf{T}$ for all models $\mathcal{I}'$ and all variables $Y_j \in \{Y_2, \ldots, Y_k\}$.
- Therefore, and because ψ is satisfiable, $\mathcal{I}'(Y_1) = \mathbf{T}$ has to hold for all models $\mathcal{I}'$.
- As per Definition of $\mathcal{I}$, we get $\mathcal{I}(Y_1) = \mathbf{T}$. Consequently, ψ is true under $\mathcal{I}$.

Exercise 2.1 (a)

Model for satisfiable Horn Formulae – Conclusion

We have shown that for an arbitrary satisfiable Horn formula ψ, all clauses are satisfied under interpretation $\mathcal{I}$.

$$\mathcal{I}(X_i) = \begin{cases} \mathbf{T} & \text{if } \mathcal{I}'(X_i) = \mathbf{T} \text{ for all models } \mathcal{I}' \text{ of } \psi \\ \mathbf{F} & \text{otherwise} \end{cases}$$

Thus, $\mathcal{I}$ is a model of ψ. $\square$

Apply (a) in order to show that there exists a formula which has no logically equivalent Horn formula.

Idea: We try to find a formula ϕ, which is satisfiable but for which the statement proven in (a) does not hold.

Exercise 2.1 (b)

Are there formulae without equivalent Horn formulae?

Consider $\phi = A \vee B$ and the two models of ϕ

$$\mathcal{I}_1 : A \rightarrow \mathbf{T}, B \rightarrow \mathbf{F}$$
$$\mathcal{I}_2 : A \rightarrow \mathbf{F}, B \rightarrow \mathbf{T}$$

Assume that there is a Horn formula ϕ' that is logically equivalent to ϕ. Thus, $\mathcal{I}_1$ and $\mathcal{I}_2$ are models of ϕ'. Already these two models (which are not the only models of ϕ') are sufficient to show that the interpretation $\mathcal{I}$ defined in (a) maps both A and B to **F**. As shown in (a), $\mathcal{I}$ is a model of ϕ', since ϕ' is a satisfiable Horn formula. But obviously $A \vee B$ is not true under $\mathcal{I}$ and, thus, ϕ' is not logically equivalent to $\phi \rightsquigarrow$ contradiction.

Exercise 2.2 (a)

Classification of Expressions

Classify the following expressions as terms, ground terms, atoms, formulae, sentences, or statements in meta language. If there is more than one possibility for an expression please list them all. The usage of symbols complies with the convention introduced with the syntax of predicate logic.

(a) $f(y)$

(b) $g(f(a), h(b, c))$

(c) $P(a, x)$

(d) $Q(a) \vee \neg Q(a)$ is valid.

(e) $P(a, b) \wedge Q(c)$

(f) $\mathcal{I}, \alpha \models P(f(x), f(a))$

(g) $Q(a) \vee P(a, b) \equiv P(b, a) \vee Q(b)$

(h) $\forall x(P(x, y) \vee Q(x))$

(i) $\forall x(\exists y(P(x, y) \wedge Q(b)) \vee P(x, y))$

(j) $\forall x(\exists y(P(x, y) \rightarrow Q(g(x, y))) \vee P(f(x), a))$

Exercise 2.2 (a)

Classification of Expressions – Terms

Terms can be variables, constants or function applications:

(a) $f(y)$

(b) $g(f(a), h(b, c))$

(c) $P(a, x)$

(d) $Q(a) \vee \neg Q(a)$ is valid.

(e) $P(a, b) \wedge Q(c)$

(f) $\mathcal{I}, \alpha \models P(f(x), f(a))$

(g) $Q(a) \vee P(a, b) \equiv P(b, a) \vee Q(b)$

(h) $\forall x (P(x, y) \vee Q(x))$

(i) $\forall x (\exists y (P(x, y) \wedge Q(b)) \vee P(x, y))$

(j) $\forall x (\exists y (P(x, y) \rightarrow Q(g(x, y))) \vee P(f(x), a))$

Exercise 2.2 (a)

Classification of Expressions – Ground Terms

Ground terms are terms without variables:

(a) $f(y)$

(b) $g(f(a), h(b, c))$

(c) $P(a, x)$

(d) $Q(a) \lor \neg Q(a)$ is valid.

(e) $P(a, b) \land Q(c)$

(f) $\mathcal{I}, \alpha \models P(f(x), f(a))$

(g) $Q(a) \lor P(a, b) \equiv P(b, a) \lor Q(b)$

(h) $\forall x (P(x, y) \lor Q(x))$

(i) $\forall x (\exists y (P(x, y) \land Q(b)) \lor P(x, y))$

(j) $\forall x (\exists y (P(x, y) \rightarrow Q(g(x, y))) \lor P(f(x), a))$

Exercise 2.2 (a)

Classification of Expressions – Formulae

Formulae can be predicates, term equivalences or propositional connectives and quantifications of other formulae:

(a) $f(y)$

(b) $g(f(a), h(b, c))$

(c) $P(a, x)$

(d) $Q(a) \lor \neg Q(a)$ is valid.

(e) $P(a, b) \land Q(c)$

(f) $\mathcal{I}, \alpha \models P(f(x), f(a))$

(g) $Q(a) \lor P(a, b) \equiv P(b, a) \lor Q(b)$

(h) $\forall x(P(x, y) \lor Q(x))$

(i) $\forall x(\exists y(P(x, y) \land Q(b)) \lor P(x, y))$

(j) $\forall x(\exists y(P(x, y) \rightarrow Q(g(x, y))) \lor P(f(x), a))$

Exercise 2.2 (a)

Classification of Expressions – Atoms

Atoms are formulae consisting of a single predicate:

(a) $f(y)$

(b) $g(f(a), h(b, c))$

(c) $P(a, x)$

(d) $Q(a) \vee \neg Q(a)$ is valid.

(e) $P(a, b) \wedge Q(c)$

(f) $\mathcal{I}, \alpha \models P(f(x), f(a))$

(g) $Q(a) \vee P(a, b) \equiv P(b, a) \vee Q(b)$

(h) $\forall x(P(x, y) \vee Q(x))$

(i) $\forall x(\exists y(P(x, y) \wedge Q(b)) \vee P(x, y))$

(j) $\forall x(\exists y(P(x, y) \rightarrow Q(g(x, y))) \vee P(f(x), a))$

Exercise 2.2 (a)

Classification of Expressions – Sentences

Sentences (or closed formulae) are formulae without free variables; all variables are quantified:

(a) $f(y)$

(b) $g(f(a), h(b,c))$

(c) $P(a,x)$

(d) $Q(a) \lor \neg Q(a)$ is valid.

(e) $P(a,b) \land Q(c)$

(f) $\mathcal{I}, \alpha \models P(f(x), f(a))$

(g) $Q(a) \lor P(a,b) \equiv P(b,a) \lor Q(b)$

(h) $\forall x(P(x,y) \lor Q(x))$

(i) $\forall x(\exists y(P(x,y) \land Q(b)) \lor P(x,y))$

(j) $\forall x(\exists y(P(x,y) \rightarrow Q(g(x,y))) \lor P(f(x), a))$

Exercise 2.2 (a)

Classification of Expressions – Statements in Meta Language

Statements in meta language are expressions that convey information outside the scope of formulae:

(a) $f(y)$

(b) $g(f(a), h(b,c))$

(c) $P(a,x)$

(d) $Q(a) \lor \neg Q(a)$ is valid.

(e) $P(a,b) \land Q(c)$

(f) $\mathcal{I}, \alpha \models P(f(x), f(a))$

(g) $Q(a) \lor P(a,b) \equiv P(b,a) \lor Q(b)$

(h) $\forall x(P(x,y) \lor Q(x))$

(i) $\forall x(\exists y(P(x,y) \land Q(b)) \lor P(x,y))$

(j) $\forall x(\exists y(P(x,y) \rightarrow Q(g(x,y))) \lor P(f(x),a))$

Exercise 2.2 (b)

Interpretation of a Theory

Consider the following theory:

$$\Theta = \left\{ \begin{array}{l} \forall x \neg P(x,x) \\ \forall x \forall y \forall z (P(x,y) \wedge P(y,z) \rightarrow P(x,z)) \\ \forall x \exists y P(x,y) \\ \neg \exists y P(y,a) \end{array} \right\}$$

Specify an interpretation $\mathcal{I} = \left\langle \mathcal{D}, \cdot^{\mathcal{I}} \right\rangle$ *with* $\mathcal{I} \models \Theta$ *(with proof). Does* Θ *have a model that is defined on a finite domain* D*?*

Exercise 2.2 (b)

Interpretation of a Theory – Examining the Theory

First we take a look at each element of theory to get an idea what an interpretation might look like.

- $\forall x \neg P(x,x)$
  The relation P has to be irreflexive: No P relates no element to itself.
- $\forall x \forall y \forall z (P(x,y) \land P(y,z) \rightarrow P(x,z)))$
  The relation P has to be transitive: If P relates x to y and y to z, then P also relates x to z.
- $\forall x \exists y P(x,y)$
  The relation P has to be serial: P relates each element to at least one element.
- $\neg \exists y P(y,a)$
  There has to be a constant a to which P relates no element to.

Idea: The relation P implies a strict ordering of elements, with one limiting constant a. This seems to fit the natural numbers along with ordering $<$ and boundary element 0.

Exercise 2.2 (b)

Interpretation of a Theory – Ordering of Natural Numbers

We choose the interpretation $\mathcal{I} = \left\langle \mathcal{D}, \cdot^{\mathcal{I}} \right\rangle$, with $\mathcal{D} := \mathbb{N}$ and $a^{\mathcal{I}} := 0, P^{\mathcal{I}} := <^{\mathbb{N}} = \left\{ (n_1, n_2) \in \mathbb{N}^2 \mid n_1 < n_2 \right\}$ where $<$ is the usual ordering of natural numbers. It now remains to show that this interpretation satisfies all formulae in Θ:

- $\mathcal{I} \models \forall x \neg P(x,x)$:
  $<^{\mathbb{N}}$ is irreflexive; no number is lesser than itself.
- $\mathcal{I} \models \forall x \forall y \forall z (P(x,y) \wedge P(y,z) \rightarrow P(x,z))$:
  $<^{\mathbb{N}}$ is transitive; if $x < y$ and $y < z$ then $x < z$.
- $\mathcal{I} \models \forall x \exists y P(x,y)$:
  $<^{\mathbb{N}}$ is serial; for each given number you can find a larger one.
- $\mathcal{I} \models \neg \exists y P(y,a)$: 0 is the smallest natural number.

With all formulae satisfied, we conclude $\mathcal{I} \models \Theta$.

Exercise 2.2 (b)

Interpretation of a Theory – Is there a model with a finite domain?

We now try to find a interpretation $\mathcal{I}' = \left\langle \mathcal{D}', \cdot^{\mathcal{I}} \right\rangle$ with $\mathcal{I}' \models \Theta$ where $\mathcal{D}'$ is a finite set $\{x_1, \ldots, x_n\}$.

- To satisfy the third sentence of Θ (seriality), each x_i has to be related to at least one element. Due to the first sentence (irreflexivity) this element cannot be x_i itself.
- Furthermore, due to the second sentence (transitivity), a chain of relations cannot be cyclic, since this would entail reflexive relations. Consequently, P has to be introduce a strict ordering of the elements. W.l.o.g we can assume the ordering $x_0 \prec x_1 \prec \ldots \prec x_n$ where $x_i \prec x_j$ stands for $P(x_i, x_j)$.
- Clearly this ordering does not satisfy seriality for the last element x_n. However, it is impossible to introduce a relation to a previous element while ensuring both irreflexivity and transitivity.

This shows that a model for Θ cannot have a finite domain.