On modal definability of Horn formulas

Stanislav Kikot

Abstract

In this short paper we give a criterion of modal definability of a first-order universal Horn sentence with exactly one positive atom in terms of its graph. As a consequence we obtain that every modal logic axiomatized by a single modal Horn formula (i.e. of the form $K + \phi$ where $\phi$ is a modal Horn formula) is Kripke complete.

Modal definability of first-order formulas has been intensively studied in modal logic, and even applied to automatic reasoning [9]. On the one hand, it has a nice Goldblatt-Thomason characterization [4], on the other hand, the problem “decide whether a first-order formula is modally definable” is in general undecidable [2]. But the cause of this undecidability is in the undecidability of first-order logic, so when we restrict attention to a fragment with decidable implication, we are likely to obtain an algorithmic criterion for modal definability, as in this paper. Also this research is motivated by scrutinizing Theorem 5.9 of [3] saying that if $L_1$ and $L_2$ are Kripke complete and Horn axiomatizable unimodal logics, then $L_1 \times L_2 = [L_1, L_2]$ and studying whether Horn axiomatizability implies Kripke completeness. We give the positive answer to the last question for the case of a single universal Horn sentence with exactly one positive atom, but in general this problem seems to be open.

Consider the classical first-order language $L_{f\Lambda}$ in the signature consisting of only binary predicates $R_\lambda$ indexed by a finite set $\Lambda$. An atom is a formula of the form $x_i R_\lambda x_j$, where $x_i$ and $x_j$ are object variables and $\lambda \in \Lambda$. Universal Horn sentences (in short, Horn formulas) are closed (i.e. without free variables) formulas of the form $\forall x_1 \ldots \forall x_n (\psi \to \phi)$, where $\psi$ is a conjunction of atoms and $\phi$ is an atom. Allowing $\lor$ in $\psi$ as in [3] is equivalent to considering conjunctions of such formulas. Universal Horn sentences can be represented by tuples of the form $D = (W^D, (R_\lambda^D : \lambda \in \Lambda), \alpha, \beta, \lambda_0)$, where $W^D = \{x_1, \ldots, x_n\}$ is a finite set, $R_\lambda^D$ are binary relations on $W^D$, $\alpha, \beta \in W^D$ and $\lambda_0 \in \Lambda$. Such a tuple $D$, called a diagram, gives rise to the Horn formula

$$E^D = \forall x_1 \ldots \forall x_n \left( \bigwedge_{x_i, x_j \in R_\lambda^D} x_i R_{\lambda} x_j \rightarrow \alpha R_{\lambda_0} \beta \right).$$

For a diagram $D$, define its size $|D| = \sum_{\lambda \in \Lambda} |R_\lambda^D|$, where $|R_\lambda^D|$ denotes the cardinality of $R_\lambda^D$. A diagram $D$ is called minimal if there is no diagram $D'$ of size less than $|D|$ such that $E^{D'} \equiv E^D$, where $\equiv$ denotes the predicate calculus equivalence. A diagram $D$ is called non-trivial if $E^D$ is not equivalent to $\top$.

We also consider the modal language $Ml\Lambda$ with countably many propositional variables $p_1, p_2, \ldots$, unary modalities $\Diamond_\lambda$ and their duals $\Box_\lambda$, where $\lambda \in \Lambda$ and boolean connectives $\land, \lor, \neg, \to$. A Kripke frame is an $L_{f\Lambda}$-structure $F = (W, (R_\lambda : \lambda \in \Lambda))$. A valuation of propositional variables in $F$ is a map $\theta$ assigning to any $p_i$ a set $\theta(p_i) \subseteq W$. A Kripke model
built on a frame $F$ is a pair $M = (F, \theta)$ where $\theta$ is a valuation of propositional variables in $F$. The truth of modal formula $\phi$ in a point $x$ of Kripke model $M$ is defined in the standard way. A modal formula $\phi$ is valid on a Kripke frame $F$ (denoted $F \models \phi$) if $\phi$ is true in every point of every model $M$ built on $F$.

An $\mathcal{L}f_\Lambda$-sentence $E$ is called modally definable if there exists a modal formula $\phi$ such that, for any Kripke frame $F$, $F \models E$ iff $F \models \phi$. Here $\vdash$ on the left-hand side means the classical truth of an $\mathcal{L}f_\Lambda$-formula in $\mathcal{L}f_\Lambda$-structure, while $\models$ on the right-hand side means the validity of a modal formula in a Kripke frame. If this equivalence holds and, in addition, $E$ is a universal Horn sentence, then $\phi$ is called a modal Horn formula.

Consider a finite $\mathcal{L}f_\Lambda$-structure $F = (W, (R_\lambda : \lambda \in \Lambda))$, where, for all $\lambda \in \Lambda$, $R_\lambda \subseteq W \times W$. A sequence $x_1, \lambda_1, x_2, \lambda_2, \ldots, x_n$, where $x_i \in W^T$, $\lambda_i \in \Lambda$ and $(x_i, x_{i+1}) \in R_{\lambda_i}$ for $1 \leq i \leq n - 1$ is called a directed path from $x_1$ to $x_n$ in $F$. The definition of an undirected path from $x_1$ to $x_n$ is obtained by replacing $(x_i, x_{i+1}) \in R_{\lambda_i}$ with $(x_i, x_{i+1}) \in R_{\lambda_i} \cup R_{\lambda_i}^{-1}$. An $\mathcal{L}f_\Lambda$-structure $F = (W, (R_\lambda : \lambda \in \Lambda))$ is called a directed tree if there is a point $r \in W$ such that the following holds:

1. $(R_\lambda)^{-1}(r) = \emptyset$ for all $\lambda \in \Lambda$,
2. for every point $x \neq r$, there exists a unique directed path from $r$ to $x$.

**THEOREM 1.** The Horn formula $E^D$ corresponding to a minimal non-trivial diagram $D = (W^D, (R^D_\lambda : \lambda \in \Lambda), \alpha, \beta, \lambda_0)$ is modally definable iff the $\mathcal{L}f_\Lambda$-structure $(W^D, (R^D_\lambda : \lambda \in \Lambda))$ is a directed tree.

The proof of the ‘if’ direction is simple: if $(W^D, (R^D_\lambda : \lambda \in \Lambda))$ is a directed tree, then all points $x_i$ except the root $x_0$ have a unique predecessor $x_{pr(i)}$ such that $x_{pr(i)}R_{\lambda(i)}x_i$ for some $\lambda(i) \in \Lambda$. Assuming that the $x$ are numbered in such a way that, for all $i$, $pr(i) < i$ and using the restricted universal quantifier

$$(\forall x_i \mathcal{D} \lambda x_j)A \equiv \forall x_i(x_jR_\lambda x_i \rightarrow A),$$

we can rewrite $E^D$ as

$$\forall x_0(\forall x_1 \mathcal{D} \lambda(1) x_0) (\forall x_2 \mathcal{D} \lambda(2) x_{pr(2)}) \ldots (\forall x_n \mathcal{D} \lambda(n) x_{pr(n)}) (\alpha R_{\lambda_0} \beta).$$

This is obviously a Kracht formula [6], [7], so it is modally definable by a Sahlqvist formula. The proof of the ‘only if’ direction follows from lemmas 2 and 4 and the fact that all modally definable properties are preserved under disjoint unions and bounded morphisms (e.g. [1]). Together with the Sahlqvist completeness theorem it gives us that any modal logic axiomatizable by a single modal Horn formula is Kripke complete. The complexity of similar logics is studied in [5].

Consider two $\mathcal{L}f_\Lambda$-structures $F^1 = (W^1, (R^1_\lambda : \lambda \in \Lambda))$ and $F^2 = (W^2, (R^2_\lambda : \lambda \in \Lambda))$. A map $g : W^1 \rightarrow W^2$ is called a homomorphism from $F^1$ to $F^2$ if, for any $\lambda \in \Lambda$ and $a, b \in W^1$, $aR^1_\lambda b$ implies $f(a)R^2_\lambda f(b)$. For a finite $\mathcal{L}f_\Lambda$-structure $F = (W, (R_\lambda : \lambda \in \Lambda))$ and a (diagram of a) Horn formula $D = (W^D, (R^D_\lambda : \lambda \in \Lambda), \alpha, \beta, \lambda_0)$ we define a Horn closure $F^*_D$ in the following way. Set $F^0_D = F$. Let $F^i_D^{-1} = (W, (R^i_D^{-1} : \lambda \in \Lambda))$ be already defined. Let $\mathcal{G}_i$ be the set of all homomorphisms from $(W^D, (R^D_\lambda : \lambda \in \Lambda))$ to $F^i_D^{-1}$. Set $F^i_D = (W, (R^i_D : \lambda \in \Lambda))$ where

$$R^i_{\lambda_0} = R^i_{\lambda_0}^{-1} \cup \bigcup_{g \in \mathcal{G}_i} \{ (g(\alpha), g(\beta)) \}$$
and \( R^i_\lambda = R^{i-1}_\lambda \) for \( \lambda \neq \lambda_0 \). Since \( W \) is finite, there exists \( n \) such that \( F_D^n = F_D^{n+1} = F_D^{n+2} \), and so on. Then we set \( F_D^* = F_D^n \) for such \( n \). This construction generalizes the well-known transitive and symmetric closure.

An \( \mathcal{L}_f A \)-structure \( F = (W, (R_\lambda : \lambda \in \Lambda)) \) and a diagram \( D = (F, \alpha, \beta, \lambda) \) are called connected if any two different points of \( W \) may be connected by an undirected way.

**LEMMA 2.** Take a minimal non-trivial diagram \( D = (W^D, (R^D_\lambda : \lambda \in \Lambda), \alpha, \beta, \lambda_0) \). If an \( \mathcal{L}_f A \)-structure \( G^D = (W^D, (R^D_\lambda : \lambda \in \Lambda)) \) is not connected then \( E^D \) is not preserved under disjoint unions.

*Proof.* First suppose that \( \alpha \) and \( \beta \) belong to different connected components of \( G^D \). Then take \( F = G^D \) and its Horn closure \( F^*_D \). Thus we have \( F^*_D \models E^D \) but \( F^*_D \sqcup F^*_D \not\models E^D \), and the lemma is proved.

Then consider the case where \( G^D \) is split into connected components \( K_1, \ldots, K_n \) and \( \alpha \) and \( \beta \) belong to the same connected component, say, \( K_1 \). Note that since \( D \) is minimal, there is no homomorphism from \( K_2 \) to \( K_1 \), otherwise we can throw \( K_2 \) out of a diagram without affecting \( E^D \) semantically. Thus we have \( K_1 \models E^D \), since there is no homomorphism from \( G^D \) in \( K_1 \) because of \( K_2 \). Put \( D' = (K_1, \alpha, \beta, \lambda_0) \). Then \( (G^D \setminus K_1)^{\mathcal{L}_f}_D \models E^D \) (since \( E^D \models E^D \)) but \( K_1 \sqcup (G^D \setminus K_1)^{\mathcal{L}_f}_D \not\models E^D \) (because of the identity homomorphism of \( G^D \) into itself and non-triviality of \( D \)). \( \Box \)

**LEMMA 3.** Consider two diagrams \( D = (W^D, (R^D_\lambda : \lambda \in \Lambda), \alpha, \beta, \lambda_0) \) and \( D' = (W'^D, (R'^D_\lambda : \lambda \in \Lambda), \alpha', \beta', \lambda'_0) \). Put \( F = (W^D, (R^D_\lambda : \lambda \in \Lambda)) \). Then \( F^*_D \models E^D \) implies \( E^{D'} \models E^D \).

*Proof.* Take any \( G = (W, (R_\lambda : \lambda \in \Lambda)) \). Assume that \( G \models E^{D'} \) and prove that \( G \models E^D \). Take a homomorphism \( h \) from \( F \) to \( G \). Now execute the process of construction of \( F^*_D \) and copy any its step by \( h \) into \( G \), applying \( G \models E^{D'} \) for each new edge. Finally we will obtain that \( h(\alpha)R_{\lambda_0}h(\beta) \). \( \Box \)

**LEMMA 4.** Let \( D \) be a minimal non-trivial diagram. Then if \( G^D = (W^D, (R^D_\lambda : \lambda \in \Lambda)) \) contains a directed cycle or a point \( c \) with two incoming arrows then \( E^D \) is not preserved under bounded morphism.

*Proof.* First suppose that \( G^D \) contains a directed cycle. Then take \( F = G^D \) and it unravelling \( F^u = ((W^D)^u, ((R^D)^u_\lambda : \lambda \in \Lambda)) \), where \( (W^D)^u \) is the set of all directed paths in \( F \), with a natural bounded morphism \( f : (W^D)^u \to W^D \); sending each path to its last point, and \( (R^D)^u_\lambda \) defined in a standard way: for \( x, y \in (W^D)^u \) \( x(R^D)^u_\lambda y \) iff \( y = x, \lambda, f(y) \). Since \( D \) is non-trivial, \( F \not\models E^D \). But \( F^u \models E^D \); since there is no homomorphism from \( G^D \) to the tree \( F^u \) because of a directed cycle in \( G^D \), so the lemma is proved.

Now assume that \( G^B \) contains a vertex with at least two incoming edges. It means that there exist points \( a, b, c \in W^D \) and \( \lambda_1, \lambda_2 \in \Lambda \) such that \( (a, c) \in R^D_{\lambda_1} \) and \( (b, c) \in R^D_{\lambda_2} \). If \( \lambda_1 \neq \lambda_2 \), the same argument as for the directed cycle works: a point with two incoming arrows of different kinds cannot be embedded into the tree \( F^u \).

But if \( a \neq b \) and \( \lambda_1 = \lambda_2 \) it may still happen that there is a homomorphism \( h \) from \( G^B \) to \( F^u \), in this case \( h(a) = h(b) \). So we consider the set \( \mathcal{T} \) of all directed trees \( T \) such that there exists a surjective homomorphism from \( G^B \) to \( T \). We claim that there exists a directed tree \( T_0 \in \mathcal{T} \) such that for all \( T \in \mathcal{T} \) there exists a surjective homomorphism from \( T_0 \) to \( T \).

Let \( \sim \) be the smallest equivalence relation on \( W^D \) satisfying condition (cf. [8])
if there exists $a, b, c, c'$ such that $aR^D \lambda c, bR^D \lambda c'$ and $c \sim c'$, then $a \sim b$.

Define $T_0 = (W^0, (R^0_\lambda : \lambda \in \Lambda))$ where $W^0 = W^D/\sim$, and for equivalence classes $A, B \in W^0$ $AR^0_\lambda B$ iff there exist $a \in A$ and $b \in B$ such that $aR_\lambda b$. In other words, $T^0$ is obtained from $G^D$ by a sequence of following reductions: if there exist $a, b, c \in W^D$ such that $aR^D_\lambda c$ and $bR^D_\lambda c$, then join $a$ and $b$ into one point. The main property of $\sim$ is that for every homomorphism $g$ from $G^D$ to a directed tree $T$ $a \sim b$ implies $g(a) = g(b)$, that is every such $g$ factors through $T_0$.

Let $h$ be the natural projection from $G^B$ to $T_0$. Consider the diagram $D' = (T_0, h(\alpha), h(\beta), \lambda_0)$. In any case, a homomorphism from $G^B$ to $T_0$ implies that $E^D \models E^D'$, and a vertex with two incoming edges in $G^B$ implies that $|D'| < |D|$. Since $D$ is minimal, $E^D' \not\models E^D$ and according to Lemma 3 it follows that $F^*_D' \not\models E^D$.

Now we can prove the lemma, since $(F^u)^*_D' \models E^D$ (use universal property of $T_0$), $F^*_D' \not\models E^D$ and $f$ is a p-morphism not only from $F^u$ to $F$, but also from $(F^u)^*_D$ to $F^*_D$.

**Acknowledgements.** This work was partially supported by the U.K. EPSRC grant EP/H05099X/1. The author also thanks V. Shehtman for pointing out this problem and his interest in progress.

**References**


