Finite Characterisations of Modal Formulas

Balder ten Cate[1*] and Raoul Koudijs[1]

ILLC, University of Amsterdam

Abstract

We initiate the study of finite characterizations and exact learnability of modal languages. A finite characterization of a modal formula w.r.t. a set of formulas is a finite set of finite models (labelled either positive or negative) which distinguishes this formula from every other formula from that set. A modal language $\mathcal{L}$ is finitely characterizable if every $\mathcal{L}$-formula has a finite characterization w.r.t. $\mathcal{L}$. This definition can be applied not only to the basic modal logic **K**, but to arbitrary normal modal logics. We show that a normal modal logic is finitely characterizable (for the full modal language) iff it is locally tabular. This shows that finite characterizations with respect to the full modal language are rare, and hence motivates the study of finite characterizations for fragments of the full modal language. Our main result is that the positive modal language without the truth-constants $\top$ and $\bot$ is finitely characterizable. Moreover, we show that this result is essentially optimal: finite characterizations no longer exist when the language is extended with the truth constant $\bot$ or with all but very limited forms of negation.

1 Introduction

We study the existence of finite characterizations of modal formulas. A finite characterization of a formula φ w.r.t. a set of formulas $\mathcal{L}$ is a finite set of finite models that distinguishes φ from every other formula in $\mathcal{L}$. Such finite characterizations are a precondition for the existence of *exact learning algorithms* for 'reverse-engineering' a hidden goal formula from examples in Angluin's model of exact learning with membership queries [1]. Our interest in exact learnability is motivated by applications in description logic. But besides learnability, the generation of examples consistent with a given formula can be used for e.g. query visualization and debugging (see e.g. [7] for a more detailed discussion of such applications). The exhaustive nature of the examples is of additional value, as they essentially display all 'ways' in which the query can be satisfied or falsified.

In this extended abstract, we only provide a high level description of our results and proof techniques. Detailed proofs can be found here: `https://bit.ly/3LCtmQt`.

2 Preliminaries

Given a set of propositional variables Prop and a set of connectives $C \subseteq \{\wedge, \vee, \Diamond, \Box, \top, \bot\}$, let $\mathcal{L}_C[\text{Prop}]$ (or simply $\mathcal{L}_C$ when Prop is clear from context) denote the collection of all modal formulas generated from literals (i.e. positive or negated propositional variables) from Prop, using the connectives in C. Note that all such formulas are in negation normal form, i.e. negations may only occur in front of propositional variables. Thus, $\mathcal{L}_{\Box,\Diamond,\wedge,\vee,\top,\bot}[\text{Prop}]$ is the set of all modal formulas with variables in Prop in negation normal form. Further, for any modal

*Supported by the European Union's Horizon 2020 research and innovation programme under grant MSCA-101031081.

fragment $\mathcal{L}$ as defined above, $\mathcal{L}^+$ and $\mathcal{L}^-$ denote the set of positive, respectively negative $\mathcal{L}$ formulas, where a formula φ is *positive* if no $p \in var(\varphi)$ occurs negated, and *negative* if all $p \in var(\varphi)$ occur only negated. We will use *modal language* to refer to any such fragment. By the *full modal language* we will mean $\mathcal{L}_{\Box,\Diamond,\wedge,\vee,\top,\bot}[\text{Prop}]$.

For a modal formula φ, let $var(\varphi)$ denote the set of variables occurring in φ and $d(\varphi)$ its *modal depth*, i.e. the nesting depth of $\Diamond$'s and $\Box$'s in φ.

A *normal modal logic* is a collection of modal formulas containing all instances of the K-axiom $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$ and closed under uniform substitution, modus ponens and generalisation.

A (Kripke) *model* is a triple $M = (dom(M), R, v)$ where $dom(M)$ is the a set of 'possible worlds', $R \subseteq dom(M) \times dom(M)$ a binary 'accessibility' relation and a valuation $V : \text{Prop} \to \mathcal{P}(W)$. A *pointed model* is a pair M, s of a Kripke model M together with a state $s \in dom(M)$. A (Kripke) *frame* is a model without its valuation.

3 Finite Characterizations

First, we define what a finite characterization means in the context of modal logic.

Definition 1. (Finite characterizations) A *finite characterization* of a formula $\varphi \in \mathcal{L}[\text{Prop}]$ w.r.t. $\mathcal{L}[\text{Prop}]$ is a pair of finite sets of finite pointed models $\mathbb{E} = (E^+, E^-)$ such that (i) φ *fits* (E^+, E^-), i.e. $E, e \models \varphi$ for all $(E, e) \in E^+$ and $E, e \not\models \varphi$ for all $(E, e) \in E^-$ and (ii) φ is the only formula in $\mathcal{L}[\text{Prop}]$ which fits (E^+, E^-), i.e. if $\psi \in \mathcal{L}[\text{Prop}]$ satisfies condition (i) then $\varphi \equiv \psi$. A modal language $\mathcal{L}$ is *finitely characterizable* if for every finite set of propositional variables Prop, every $\varphi \in \mathcal{L}[\text{Prop}]$ has a finite characterization w.r.t. $\mathcal{L}[\text{Prop}]$.

Thus if (E^+, E^-) is a finite characterization of a formula $\varphi \in \mathcal{L}[\text{Prop}]$ w.r.t. $\mathcal{L}[\text{Prop}]$, then for every $\psi \in \mathcal{L}[\text{Prop}]$ with $\varphi \not\equiv \psi$, E^+ contains a finite model of $\varphi \wedge \neg\psi$ or E^- contains a finite model of $\neg\varphi \wedge \psi$.

For example, the formula $p \wedge q$ has a finite characterization w.r.t. $\mathcal{L}^+_\wedge[\text{Prop}]$ with $\text{Prop} = \{p, q, r\}$, namely $(\{\cdot_{p,q}\}, \{\cdot_p, \cdot_q\})$, where "$\cdot_P$" is the single point model where all $p \in P$ are true.

Our motivation for studying finite characterizations, comes from *computational learning theory*. Specifically, finite characterizability is a necessary precondition for *exact learnability with membership queries* in Dana Angluin's interactive model of exact learning [1]. In our context, exact learnability with membership corresponds to a setting in which the learner has to identify a formula by asking question to an oracle, where each question is of the form "is the formula true or false in pointed model (M, w)?" This can also be viewed as a 'reverse engineering' task, where a formula has to be identified based on its behaviour on only a finite set of models. Exact learnability has recently gained a renewed interest in the description logic literature. We comment more on the connection with description logic in Section 4.

Our starting observation is:

Theorem 1. *The full modal language is not finitely characterizable.*

Proof. It suffices to give one counterexample, so suppose that e.g. $\varphi = \Box\bot$ had a finite characterization (E^+, E^-) w.r.t. the full modal language. Observe that for each n, $M, s \models \Box^{n+1}\bot \wedge \Diamond^n\top$ iff $height(M, s) = n$, where the *height* of a pointed model M, s is the length of the longest path in M starting at s, or ∞ if there is no finite upper bound. Every finite characterization can only contain pointed models up to some bounded height $< n$ (by choice of n) or must contain some model of infinite (∞) height. In either case, it follows that no negative example $(E,e) \in E^-$ satisfies $\Box^{n+1}\bot \wedge \Diamond^n\top$. Hence for large enough n, $\varphi \vee (\Box^{n+1}\bot \wedge \Diamond^n\top)$ also fits (E^+, E^-), yet is clearly not equivalent to φ. $\Box$

In fact, by a variation of the same argument, we can show that *no* modal formula has a finite characterization w.r.t. the full modal language. Theorem 1 raises two questions, namely: *do finite characterizations exist in other modal logics than* **K**, and *which fragments of modal logic admit finite characterizations.* We address each of these two questions next.

We first generalize Definition 1 as follows (whereby Theorem 1 becomes a result about the special case of the basic normal modal logic **K**): a finite characterization of a modal formula φ with $var(\varphi) \subseteq \mathrm{Prop}$ w.r.t. a normal modal logic L is a finite set (E^+, E^-) of finite pointed models based on L frames such that (i) φ fits (E^+, E^-) and (ii) if ψ with $var(\psi) \subseteq \mathrm{Prop}$ fits (E^+, E^-) then $\varphi \equiv_L \psi$, where $\varphi \equiv_L \psi$ iff $\varphi \leftrightarrow \psi \in L$. We say that a normal modal logic L is finitely characterizable if for every finite set Prop, every modal φ with $var(\varphi) \subseteq \mathrm{Prop}$ has a finite characterization w.r.t. L. We can give a complete characterization over which modally definable frame classes the full modal language is finitely characterizable.

It turns out that only very few normal modal logics are uniquely characterizable. A normal modal logic L is *locally tabular* if for every finite set Prop of propositional variables, there are only finitely many formulas φ with $var(\varphi) \subseteq \mathrm{Prop}$ up to L-equivalence.

Theorem 2. *A normal modal logic L is finitely characterizable iff it is locally tabular.*

In other words the full modal language is only finitely characterizable in the degenerate case where there are only finitely many formulas to distinguish from (up to equivalence). This result motivates the investigation of finite characterizability for modal fragments. Specifically, inspired by previous work on finite characterizability of the positive existential fragment of first order logic [2], we consider positive fragments of the full modal language.

Note that, in the remainder of this section, we only consider again the modal logic **K**.

The proof of Theorem 1 can easily be modified to show the following:[1]

Theorem 3. *$\mathcal{L}^+_{\Box,\Diamond,\wedge,\vee,\bot}$ is not finitely characterizable.*

On the other hand, based on results in [2], we can show that:

Theorem 4 (From [2])**.** *$\mathcal{L}^+_{\Diamond,\wedge}$ is finitely characterizable. Indeed, given a formula in $\mathcal{L}^+_{\Diamond,\wedge}$, we can construct a finite characterization in polynomial time.*

More precisely, it was shown in [2] that finite characterizations can be constructed in polynomial time for "c-acyclic conjunctive queries", a fragment of first-order logic that includes the standard translations of $\mathcal{L}^+_{\Diamond,\wedge}$-formulas.

Our main result here extends Theorem 4 by showing that $\mathcal{L}^+_{\Box,\Diamond,\wedge,\vee}$ is finitely characterizable.

Theorem 5. *$\mathcal{L}^+_{\Box,\Diamond,\wedge,\vee}$ is finitely characterizable.*

Theorem 3 above shows that this is essentially optimal; we leave open the question whether the fragment without $\bot$ but with $\top$ is finitely characterizable.

In the rest of this section, we outline the ideas behind the proof of Theorem 5. A key ingredient is the novel notion of *weak simulation*, which we obtain by weakening the back and forth clauses of the *simulations* studied in [3]. Simulations are themselves a weakening of bisimulations. It was shown in [3] that $\mathcal{L}^+_{\Box,\Diamond,\wedge,\vee,\top,\bot}$ is characterized by preservation under simulations.

[1] It suffices to replace $\top$ by a fresh propositional variable q in the proof of Theorem 1.

A *weak simulation* between two pointed models $(M, s), (M', s')$ is a relation $Z \subseteq M \times M'$ such that for all $(t, t') \in Z$:

(atom)	$M, s \models p$ implies $M', s' \models p$
(forth′)	If $R^M tu$, either $M, u \underline{\leftrightarrow} \circlearrowright_\emptyset$ or there is a u' with $R^{M'} t'u'$ and $(u, u') \in Z$
(back′)	If $R^{M'} t'u'$, either $M', u' \underline{\leftrightarrow} \circlearrowright_{\text{Prop}}$ or there is a u with $R^M tu$ and $(u, u') \in Z$

where $\circlearrowright_\emptyset$ denotes the single reflexive point with empty valuation, $\circlearrowright_{\text{Prop}}$ denotes the single reflexive point with full valuation and $\underline{\leftrightarrow}$ denotes bisimulation. If such Z exists, we say that M', s' *weakly simulates* M, s. The crucial weakening is witnessed by the fact that the deadlock model, i.e. the single point with no successors, weakly simulates $\circlearrowright_\emptyset$, but does not simulate it.

Because weak simulations are closed under relational composition, which is associative, the collection of pointed models and weak simulations forms a category with $\circlearrowright_\emptyset$ and $\circlearrowright_{\text{Prop}}$ as weak initial and final objects, respectively.

Theorem 6. *$\mathcal{L}^+_{\Box,\Diamond,\wedge,\vee}$ is preserved under weak simulations.*

In high level terms, the proof of Theorem 5 proceeds as follows: given a formula $\varphi \in \mathcal{L}^+_{\Box,\Diamond,\wedge,\vee}$, we show how to construct positive and negative examples $(E^+_\varphi, E^-_\varphi)$ that φ fits and which form a *duality* (a generalisation of the notion of *splittings* in lattice theory [6]) in the category of pointed models and weak simulations. By this, we mean that every pointed model either weakly simulates some positive example in E^+ or is weakly simulated by some negative example in E^-. More specifically, we show that every model of φ weakly simulates some positive example in E^+ and that every non-model of φ is weakly simulated by some negative example in E^-. It follows by Theorem 6 that any $\mathcal{L}^+_{\Box,\Diamond,\wedge,\vee}$-formula that fits E^+ is implied by φ, while every formula that fits E^- implies φ. Combined, this shows that $(E^+_\varphi, E^-_\varphi)$ is a finite characterization of φ w.r.t. $\mathcal{L}^+_{\Box,\Diamond,\wedge,\vee}$.

This proof technique was inspired by results in [7], which established a similar connection between finite characterizations for GAV schema mappings (or, equivalently, unions of conjunctive queries) and dualities in the category of finite structures and homomorphisms.

See `https://bit.ly/3LCtmQt` for more details and further results.

4 Discussion

Our construction, although effective, is non-elementary. For this reason, we cannot obtain from it an efficient exact learning algorithm. On the other hand, it follows from the results in [2] that $\mathcal{L}^+_{\Diamond,\wedge}$-formulas are polynomial-time exactly learnable with membership queries. We leave it as future work to prove matching lower bounds for our construction, and to understand more precisely which modal fragments admit polynomial-sized finite characterizations and/or are polynomial-time exactly learnable with membership queries.

Variants of Theorem 5 can be obtained for $\mathcal{L}^-_{\Box,\Diamond,\wedge,\vee}$ and, more generally, for *uniform* modal formulas, where certain propositional variables only occur positive and others only negatively.

As we mentioned, our immediate motivation for this work came from a renewed interest in exact learnability in description logic. In particular, in [2], exact learnability with membership queries is studied for the description logic $\mathcal{ELI}$. These results are extended to results on learning $\mathcal{ELI}$ concepts under DL-Lite ontologies (i.e. background theory) [4] and temporal instance queries formulated in fragments of linear time logic LTL [5]. We expect that our proof of Theorem 5 can be lifted to the poly-modal case without major changes, with implications for some description logics under the closed-world assumption.

References

[1] D. Angluin. Queries and concept learning. Mach. Learn., 2(4):319–342, 1988.

[2] B. ten Cate and V. Dalmau. Conjunctive Queries: Unique Characterizations and Exact Learnability. In Proceedings of ICDT 2021, pages 9:1–9:24, 2021.

[3] N. Kurtonina and M. D. Rijke. Simulating without negation. Journal of Logic and Computation, 7(4):501–522, 08 1997.

[4] M. Funk, J. C. Jung, and C. Lutz. Actively learning concepts and conjunctive queries under ELr-ontologies. In IJCAI, 2021.

[5] M. Fortin, B. Konev, V. Ryzhikov, Y. Savateev, F. Wolter, and M. Zakharyaschev. Unique characterisability and learnability of temporal instance queries, 2022.

[6] R. McKenzie. Equational bases and nonmodular lattice varieties. Transactions of the American Mathematical Society, 174:1–43, 1972.

[7] B. Alexe, B. t. Cate, P. G. Kolaitis, and W.-C. Tan. Characterizing schema mappings via data examples. ACM Trans. Database Syst., 36(4):23:1–23:48, Dec. 2011.