Zero-One Laws for Modal Logic

Joseph Y. Halpern
IBM Almaden Research Center
San Jose, CA 95120
halpern@almaden.ibm.com

Bruce M. Kapron
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213
bmkapron@cs.cmu.edu

Abstract

We show that a 0-1 law holds for propositional modal logic, both for structure validity and frame validity. In the case of structure validity, the result follows easily from the well-known 0-1 law for first-order logic. However, our proof gives considerably more information. It leads to an elegant axiomatization for almost-sure structure validity, and sharper complexity bounds. Since frame validity can be reduced to a Π_1^1 formula, the 0-1 law for frame validity helps delineate when 0-1 laws exist for second-order logics.

1 Introduction

Glebskii *et al.* [GKLT69] and, independently, Fagin [Fag76] established a fascinating connection between asymptotic probability and logical definability. They showed that for every property P expressible in first-order logic without function symbols is either almost-surely true or almost-surely false in finite structures: more precisely, if $\mu_n(P)$ is the fraction of structures with domain $\{1, \ldots, n\}$ in which P is true, then $\mu(P) = \lim_{n\to\infty} \mu_n(P)$ is either 0 or 1. This remarkable property is known as a *0-1 law*. Grandjean [Gra83] showed that the problem of deciding whether a formula is almost-surely true is PSPACE-complete for bounded underlying vocabularies, that is, vocabularies with an *a priori* bound on the arities of the predicates in the vocabulary. This forms a sharp contrast to Trakhtenbrot's classical theorem that says that the set of formulas which are valid in all finite structures is co-r.e., assuming that the vocabulary contains at least one binary predicate.

It is well known that first-order logic has rather limited expressive power (cf. [AU79, Fag75]). Thus, researchers have investigated asymptotic probabilities in logical languages that go beyond first-order. (See [Com88] for an overview and references.) Of most interest to us here are results for existential second-order logic (Σ_1^1). The interest in Σ_1^1 stems in part from a result of Fagin [Fag74] showing that a property is Σ_1^1 definable if and only if it is NP-computable. Full Σ_1^1 does not have a 0-1 law [Fag76]; in fact, neither does monadic Σ_1^1, where the existential quantification is over monadic predicates [Kau87]. A 0-1 law does hold for certain fragments of existential second-order logic, whose structure is characterized by the pattern of first-order quantifiers. If Ψ is a class of first-order formulas, let $\Sigma_1^1(\Psi)$ consist of the set of Σ_1^1 sentences where the first-order part is in Ψ. Kolaitis and Vardi considered first-order formulas Ψ defined by their quantifier prefix. There are exactly two *prefix classes* of first-order formulas with equality for which the validity problem is decidable [DG79]: the *Bernays-Schönfinkel class*, consisting of formulas with quantifier prefixes of the form $\exists^*\forall^*$ (i.e., a possibly empty string of existential quantifiers followed by a possibly empty string of universal quantifiers) and the *Ackerman class*, consisting of formulas with quantifier prefix $\exists^*\forall\exists^*$. The results of [KV87, KV90, PS89] show that a 0-1 law holds for $\Sigma_1^1(\Psi)$, where Ψ is the set of first-order formulas defined by some prefix class, iff Ψ is either the Bernays-Schönfinkel class or the Ackermann class. In addition, Kolaitis and Vardi show that if Ψ is either the Bernays-Schönfinkel class or the Ackermann class, the problem of deciding whether a formula in $\Sigma_1^1(\Psi)$ is almost-surely true is NEXPTIME-complete if we restrict to bounded vocabularies [KV87, KV90].

We focus here on 0-1 laws for (propositional) modal logic. Modal logic is a natural logic to investigate, given the attention it has received in the philosophical literature, and the fact that various modal logics have been shown to be of great relevance to computer science, including temporal logic [MP81] and epistemic logic (i.e., reasoning about knowledge) [Hal87]. In order to explain our results, we briefly review the syntax and semantics of modal logic. The syntax is quite simple: we start with primitive propositions and close off under negation, conjunction, and application of the modal operator $\Box$. The standard semantics for modal logic is *possible-worlds* semantics. A *frame* F is a pair $(S, \mathcal{R})$, consisting of a set of possible worlds (or states) and a binary relation $\mathcal{R}$ (called the *possibility relation*) on S. A *(Kripke) structure* M is a tuple $(S, \mathcal{R}, \pi)$ consisting of a frame $(S, \mathcal{R})$ and a truth assignment π, which assigns a truth value to each primitive proposition in each state in S. We say that the structure $(S, \mathcal{R}, \pi)$ is *based on* the frame $(S, \mathcal{R})$. $(M, s) \models \varphi$ (φ is true in state s of structure M) is defined in a straightforward way by induction on the structure of φ. (We use the relation $\mathcal{R}$ to define the semantics of $\Box$ formulas). Modal logicians have historically been interested in properties of both frames and structures [HC84, Ben83]. A formula φ is said to be *valid in structure* $M = (S, \mathcal{R}, \pi)$ if $(M, s) \models \varphi$ for every state $s \in S$. We say φ is *valid in frame* F if φ is valid in every structure M based on F. Finally, we say that φ is *structure* (resp., *frame*) *valid* if φ is valid in every structure (resp., frame). It is immediate from the definitions that a formula is structure valid if and only if it is frame valid. It is also well-known that deciding structure/frame validity is PSPACE-complete for K and S4, and co-NP-complete for S5 [Lad77].

When it comes to 0-1 laws, there are two questions we can investigate: almost-sure structure validity and almost-sure frame validity. That is, we can consider all structures with state space $\{1, \ldots, n\}$ and consider in what fraction of them a formula φ is valid, or we can consider all frames with state space $\{1, \ldots, n\}$ and ask in what fraction of them φ is valid. In both cases a 0-1 law holds: that is, a formula is valid in almost all structures (resp., frames) with state space $\{1, \ldots, n\}$ or almost none of them. However, although structure validity and frame validity coincide, almost-sure structure validity and almost-sure frame validity do not.

There is a well-known translation [Ben83] from a modal logic formula φ to a first-order logic formula φ^{fo} with the property that the fraction of structures with state space $\{1, \ldots, n\}$ for which φ is valid is precisely the same as the fraction of relational structues with domain $\{1, \ldots, n\}$ for which φ^{fo} is true. For future reference, we note that φ^{fo} has one unary predicate P corresponding to each primitive proposition p in φ, and one binary predicate R corresponding to the possibility relation $\mathcal{R}$. The 0-1 law for structure validity thus follows immediately from the 0-1 law for first-order logic in light of this translation. Our proof of the 0-1 law for structure validity does not proceed via this translation; instead, it uses the relatively straightforward observation that if φ is a consistent propositional formula, then $\Diamond\varphi$ is almost-surely structure valid (where $\Diamond$ is the dual of $\Box$). Using this observation, we can show that deciding if a formula is almost-surely structure valid is in Δ_2^p ($\Delta_2^P =_{\rm def} P^{NP}$, i.e., P with an NP oracle, and is in the second level of the polynomial time hierarchy [Sto77]). This suggests that almost-sure modal validity is easier than both modal validity and the problem of deciding if an arbitrary first-order formula is almost-surely true. Finally, we show that we can axiomatize the set of formulas that are almost-surely structure valid in a straightforward way.

The proof of the 0-1 law in the case of frame validity is far more difficult. An argument analagous to that used in the case of structure validity shows that for any modal formula φ which uses the propositional letters $p_1, \ldots, p_k$, the fraction of frames with state space $\{1, \ldots, n\}$ for which φ is valid is precisely the same as the fraction of relational structures with state space $\{1, \ldots, n\}$ for which the Π_1^1 formula $\forall P_1 \ldots P_k \varphi^{fo}$ is true. Recasting this in terms of satisfiability, the fraction of frames with state space $\{1, \ldots, n\}$ for which φ is satisfiable is precisely the same as the fraction of relational structures with domain $\{1, \ldots, n\}$ in which the Σ_1^1 formula $\exists P_1 \ldots P_k \varphi^{fo}$ is true. Let MDL be the set of first-order formulas that arise as the translation of modal formulas. Our results can thus be interpreted as showing that the class $\Sigma_1^1(MDL)$ has a 0-1 law. It is easy to show that MDL is incomparable in expressive power to both the Ackermann class and the Bernays-Schönfinkel class. We conjecture that $\Sigma_1^1(MDL)$ is incomparable in expressive power to both Σ_1^1(Ackermann) and Σ_1^1(Bernays-Schönfinkel), although we have not proved this. The techniques that we use to prove that a 0-1 law holds for frame validity involve rather delicate combinatorial arguments, and are quite different from those used by Kolaitis and Vardi. Note that $\Sigma_1^1(MDL)$ is actually a fragment of monadic Σ_1^1. Thus, our results help delineate when 0-1 laws exist for second-order logics.

We also show that the problem of deciding whether a formula is almost-surely frame valid is hard for deterministic exponential time, and thus is harder than the frame validity problem (assuming PSPACE $\neq$ EXPTIME). Notice that the vocabulary for MDL is bounded, since it involves only unary and binary predicates. To the best of our knowledge, this is the first time that deciding if a formula is almost-surely valid with respect to a class of structures has been shown to be *harder* than showing it is valid with respect to that class.

2 Preliminaries

As we mentioned in the introduction, the formulas of propositional modal logic are those obtained by starting with primitive propositions in some set Φ and closing off under negation, conjunction, and application of the modal operator $\Box$. We call the resulting language $\mathcal{L}(\Phi)$. We give semantics to these formulas via Kripke structures. A *Kripke structure M over Φ* is a tuple $(S, \mathcal{R}, \pi)$, as defined in the introduction. We use π to give the semantics for primitive propositions in Φ; the semantics of the Boolean connectives is as in propositional logic; finally, we define $\Box\varphi$ to be true at a state s if φ is true in all worlds reachable from s via the $\mathcal{R}$ relation. Thus, we have

- $(M, s) \models p$ for a $p \in \Phi$ iff $\pi(s)(p) = \mathbf{true}$.
- $(M, s) \models \neg\varphi$ iff $(M, s) \not\models \varphi$
- $(M, s) \models \varphi \wedge \psi$ iff $(M, s) \models \varphi$ and $(M, s) \models \psi$.
- $(M, s) \models \Box\varphi$ iff $(M, t) \models \varphi$ for all t such that $(s, t) \in \mathcal{R}$.

As usual, we say a formula φ is *valid* (resp., *satisfiable*) in model $M = (S, \mathcal{R}, \pi)$ if $M, s \models \varphi$ for all (resp., some) $s \in S$. We write $M \models \varphi$ if φ is valid in M We say that φ is *structure valid* if it is valid in all structures, and *structure satisfiable* if it is satisfiable in some structure. A formula φ is valid in frame F if it is valid in all models based on F; φ is satisfiable in F if it is satisfiable in some model based on F. Finally, we say φ is *frame valid* if it is valid in all frames, and *frame satisfiable* if it is satisfiable in some frame.

The logic just defined, known as K, can be axiomatized as follows [HC78]:

A1. All instances of tautologies of propositional calculus

A2. $(\Box\varphi \wedge \Box(\varphi \Rightarrow \psi)) \Rightarrow \Box\psi$

R1. From φ and $\varphi \Rightarrow \psi$ infer ψ (Modus ponens)

R2. From φ infer $\Box\varphi$ (Generalization)

Modal logicians have considered numerous modal logics other than K. The ones of most interest to us here are those that have been called T, S4, and S5. All these logics satisfy the axioms of K. T is characterized by the axioms of K together with

A3. $\Box\varphi \Rightarrow \varphi$

S4 is characterized by the axioms of T together with:

A4. $\Box\varphi \Rightarrow \Box\Box\varphi$

Finally, S5 is characterized by the axioms of S4 together with:

A5. $\neg\Box\varphi \Rightarrow \Box\neg\Box\varphi$

Let $\mathcal{M}$ consist of all Kripke structures, and let $\mathcal{M}^r$ (resp., $\mathcal{M}^{rt}$; $\mathcal{M}^{rst}$) consist of all Kripke structures where the $\mathcal{R}$ relation is reflexive (resp., reflexive and transitive; reflexive, symmetric, and transitive, i.e., an equivalence relation).

Theorem 2.1: [HC78] *K (resp., T; S4; S5) is sound and complete with respect to the structures in $\mathcal{M}$ (resp., $\mathcal{M}^r$; $\mathcal{M}^{rt}$; $\mathcal{M}^{rst}$).*

We can similarly define $\mathcal{F}$ to consist of all frames, and $\mathcal{F}^r$ (resp., $\mathcal{F}^{rt}$; $\mathcal{F}^{rst}$) to consist of all frames where the $\mathcal{R}$ relation is reflexive (resp., reflexive and transitive; and equivalence relation). It is trivial to check that structure validity and frame validity coincide, thus K (resp., T; S4; S5) is also sound and complete with respect to the frames in $\mathcal{F}$ (resp., $\mathcal{F}^r$; $\mathcal{F}^{rt}$; $\mathcal{F}^{rst}$).

Let Φ be a set of primitive propositions, and let $\mathcal{M}_{n,\Phi}$ (resp., $\mathcal{F}_{n,\Phi}$) be the set of Kripke structures (resp., frames) over Φ with state space $\{1, \ldots, n\}$. Notice that $\mathcal{M}_{n,\Phi}$ and $\mathcal{F}_{n,\Phi}$ are finite if Φ is finite. If Φ is finite, then we take $\nu_{n,\Phi}$ to be the uniform probability distribution on $\mathcal{M}_{n,\Phi}$, and take $\mu_{n,\Phi}$ to be the uniform probability distribution on $\mathcal{F}_{n,\Phi}$. Although the main interest in 0-1 laws has been for finite structures, for technical reasons, we also allow Φ to be infinite. There are a number of ways to proceed in this case, all of which turn out to be equivalent for our purposes (see [GHK92] for further discussion of this issue). If Φ is infinite, we consider the the σ-algebra over $\mathcal{M}_{n,\Phi}$ generated by $\mathcal{M}_{n,\Phi'}$ for all finite subsets Φ' of Φ. That is, given a structure $M \in \mathcal{M}_{n,\Phi'}$, we consider the subset A_M of $\mathcal{M}_{n,\Phi'}$ consisting of structures that agree with M on the propositions in Φ', and consider the σ-algebra generated by all sets of the form A_M. We define $\mu_{n,\Phi}$ on this σ-algebra so that $\mu_{n,\Phi}(A_M) = \mu_{n,\Phi'}(A_M)$. It is easy to check that this is a well-defined measure. We similarly define a $\nu_{n,\Phi}$ on $\mathcal{F}_{n,\Phi}$. For a formula $\varphi \in \mathcal{L}(\Phi)$, we write $\nu_{n,\Phi}(\varphi)$ as an abbreviation for $\nu_{n,\Phi}(\{M \in \mathcal{M}_{n,\Phi} : M \models \varphi\})$, and similarly for $\mu_n(\varphi)$. It is easy to see that if $\varphi \in \mathcal{L}(\Phi)$ and $\Phi' \supset \Phi$, then $\mu_{n,\Phi}(\varphi) = \mu_{n,\Phi'}(\varphi)$, and similarly $\nu_{n,\Phi}(\varphi) = \nu_{n,\Phi'}(\varphi)$. Thus, without loss of generality, we need to consider only finite sets Φ when computing asymptotic probabilities. We omit the subscript Φ in the rest of the paper if its role is unimportant.

Although we are mainly interested in computing $\nu(\varphi) = \lim_{n\to\infty} \nu_n(\varphi)$ and $\mu(\varphi) = \lim_{n\to\infty} \mu_n(\varphi)$, in the process we need to apply μ and ν to events other than those defined by formulas. We are also interested in computing the asymptotic limits when we restrict to structures in $\mathcal{M}^r$, $\mathcal{M}^{rt}$, and $\mathcal{M}^{rst}$, and similarly for frames. We can make the obvious analogous definitions, for example, taking $\mathcal{F}_n^r$ to be the set of frames with state space $\{1,\ldots,n\}$ in which the $\mathcal{R}$ relation is reflexive, defining $\mu_n^r(\varphi)$ to be the fraction of frames in $\mathcal{F}_n^r$ in which φ is valid, and defining $\mu^r(\varphi) = \lim_{n\to\infty} \mu_n^r(\varphi)$.

3 The expressive power of modal formulas

We begin this section by reviewing the translation from modal logic to first-order logic mentioned in the introduction. We then show that $\Sigma_1^1(MDL)$ can capture some NP-complete properties, in particular satisfiability of propositional logic.

Suppose the primitive propositions in the modal language are $p_1, p_2, \ldots$. Consider the first-order vocabulary Φ consisting of the unary predicates $P_1, P_2, \ldots$ and the binary predicate R. We now show how to translate a modal formula φ to a first-order formula $\varphi^{fo}(x)$ with one free variable x over the vocabulary Φ. We proceed by induction on structure:

- $p^{fo} = P(x)$ if p is a primitive proposition
- $(\varphi \wedge \psi)^{fo} = \varphi^{fo} \wedge \psi^{fo}$
- $(\neg\varphi)^{fo} = \neg(\varphi^{fo})$
- $(\Box\varphi)^{fo} = \forall y(R(x,y) \Rightarrow \varphi^{fo}[x/y])$,

where $\varphi^{fo}[x/y]$ is the result of replacing all free occurences of x in φ^{fo} by y. Let *MDL* be the set of first-order formulas that are of the form φ^{fo} for some modal formula φ. It is easy to see that when formulas in *MDL* are put into prenex form, we can have arbitrarily deep alternation of quantifiers. Thus, *MDL* is syntactically distinct from the Ackermann class and the Bernays-Schönfinkel class. Although we do not go into details here, we remark that the results of van Benthem [Ben83] characterizing *MDL* show that in fact it is inequivalent in expressive power to both of these classes.

Given a Kripke structure $M = (S, \mathcal{R}, \pi)$, let M^{fo} be the relational structure over Φ with domain S, where the interpretation of P_i is the set of states in S where p_i is true according to π, and the interpretation of R is $\mathcal{R}$.

Proposition 3.1: [Ben83] *φ is valid in M iff $\forall x \varphi^{fo}$ is true in M^{fo}.*

We take $\Pi_1^1(MDL)$ (resp., $\Sigma_1^1(MDL)$) to consist of formulas of the form $\forall\mathbf{P}\,\varphi$ (resp., $\exists\mathbf{P}\,\varphi$), where φ is a first-order formula in *MDL* with unary predicates in $\mathbf{P}$ and binary predicate R. Given a frame $F = (S, \mathcal{R})$, let F^{fo} be the relational structure over R with domain S, where the interpretation of R is $\mathcal{R}$. It immediately follows from Proposition 3.1 that

Proposition 3.2: *φ is valid in F iff $\forall\mathbf{P}\forall x\varphi^{fo}$ is true in F^{fo}, where $\mathbf{P}$ includes the unary predicates that appear in φ^{fo}.*

Thus, frame validity can be expressed by formulas in $\Pi_1^1(MDL)$; analogously, frame satisfiability can be expressed by formulas in $\Sigma_1^1(MDL)$. This leads us to consider the expressive power of $\Sigma_1^1(MDL)$. It is well-known that $\Sigma_1^1(MDL)$ is incomparable in expressive to first-order logic [Ben83]. We conjecture that $\Sigma_1^1(MDL)$ is incomparable in expressive power to both Σ_1^1(Ackermann) and Σ_1^1(Schönfinkel), but have no proof of this. Of more interest to us here is that, just like Σ_1^1(Schönfinkel) [KV87] and Σ_1^1(Ackermann) [KV90], $\Sigma_1^1(MDL)$ can express NP-complete properties. In particular, we now give a construction (due to Moshe Vardi) showing that $\Sigma_1^1(MDL)$ can express satisfiability of CNF formulas.

With every propositional formula α in CNF, we construct a frame F_α; we then define a modal formula φ_{SAT} such that φ_{SAT} is satisfiable in F_α iff α is a satisfiable propositional formula. This gives us the result we want.

Given α, we think of F_α as a rooted dag. From the root, we construct one successor for each clause in α. We also have a leaf node for each primitive proposition that appears in α. Suppose that β is one of the clauses in α. If the primitive proposition p is one the disjuncts that appears in β, then there is a path of length one to the node representing p. If $\neg p$ is one the disjuncts in β, then there is a path of length two from the node representing β to the node representing p. Finally, we add a path of length 3 starting at the root. (This will allow us to distinguish the root from all other nodes in F_α, none of which are at the beginning of paths of length longer than 2.) This completes the description of F_α.

We take φ_{SAT} to be $\Diamond\Diamond\Diamond \mathit{true} \wedge \Box(\Diamond q \vee \Diamond\Diamond\neg q)$.

Theorem 3.3: *If α is a propositional formula in CNF, then φ_{SAT} is satisfiable in F_α iff the CNF formula α is satisfiable.*

Proof: Suppose that α is satisfiable. Let v be a truth assignment satisfying α. Consider the structure $M = (S_\alpha, \mathcal{R}_\alpha, \pi_\alpha)$ based on F_α, where π_α is defined so that, for each primitive proposition p that appears in α, if $s \in S_\alpha$ is a leaf node representing p, then $\pi_\alpha(s)(q) = v(p)$. That is, q is true at the state s according to π_α iff p is true according to v. The truth value that π_α assigns to q at non-leaf nodes is irrelevant. It is easy to check that, if s_0 is the root of F_α, then $(M_\alpha, s_0) \models \varphi_{SAT}$; thus φ_{SAT} is satisfiable in F_α. The proof of the converse is similar, and is left to the full paper. ∎

This shows that satisfiability of CNF formulas is expressible in $\Sigma^1_1(MDL)$.

4 0-1 laws for structure validity

It is easy to see that the mapping $M \rightarrow M^{fo}$ gives a one-to-one correspondence between Kripke structures with state space $\{1, \ldots, n\}$ and relational structures over Φ with domain $\{1, \ldots, n\}$. Thus, the following corollary to Theorem 3.1 is immediate.

Corollary 4.1: *The fraction of Kripke structures with state space $\{1, \ldots, n\}$ for which φ is valid is the same as the fraction of relational structures over Φ for which φ^{fo} is true.*

Putting this together with the 0-1 law for first-order logic, we get

Corollary 4.2: *There is a 0-1 law for structure validity.*

This translation does not give us the other results claimed in our introduction. These all follow from the following simple observation. Let $\Diamond\varphi$ be an abbreviation for $\neg\Box\neg\varphi$. Thus, $(M, s) \models \Diamond\varphi$ if there is some state t such that $(s, t) \in \mathcal{R}$ and $(M, t) \models \varphi$.

Proposition 4.3: *If φ is a consistent propositional formula, then $\Diamond\varphi$ is valid in almost all structures.*

Proof: Suppose that φ mentions k primitive propositions. There are 2^k possible assignments of truth values to these primitive propositions. Since φ is consistent, then at least one of these truth assignments makes φ true. The probability that a given state in a structure with state space $\{1, \ldots, n\}$ does not have any $\mathcal{R}$-successors where φ is true is thus at most $(1 - 1/2^k)^n$. Thus, the probability that some state in such a structure does not have any $\mathcal{R}$ successors where φ is true is $\alpha(n) = 1 - (1 - (1 - 1/2^k)^n)^n$. Let β be an abbreviation for $1 - 1/2^k$. Using the binomial expansion, we have $(1 - \beta^n)^n = 1 - n\beta^n + O(n^2\beta^{2n})$. Thus, $\alpha(n) = n\beta^n + O(n^2\beta^{2n})$. It follows that $\lim_{n\rightarrow\infty} \alpha = 0$. Thus, for almost all structures, $\Diamond\varphi$ is valid. ∎

We now provide a translation from an arbitrary modal formula φ to a propositional formula φ^r, with the property that $\varphi \Leftrightarrow \varphi^r$ is almost-surely valid. We proceed by induction on the structure of formulas:

- $p^r = p$ for a primitive proposition p
- $(\varphi \wedge \psi)^r = \varphi^r \wedge \psi^r$
- $(\neg\varphi)^r = \neg\varphi^r$
- $(\Box\varphi)^r = \begin{cases} \mathit{true} & \text{if } \varphi^r \text{ is valid} \\ \mathit{false} & \text{otherwise.} \end{cases}$

Proposition 4.4: *The formula $\varphi \Leftrightarrow \varphi^r$ is valid in almost all structures.*

Proof: By a straightforward induction on the structure of φ. We leave details to the full paper. ∎

We now immediately get:

Theorem 4.5: *For all modal formulas φ, we have $\nu(\varphi) = 1$ iff the propositional formula φ^r is valid; otherwise $\nu(\varphi) = 0$.*

We next consider the complexity of computing whether a formula is almost-surely valid. The situation is surprisingly subtle. For one thing, it turns out to matter if we take Φ to be finite or infinite. Notice that if we take Φ to be finite, the complexity of computing satisfiability for propositional formulas over Φ is linear time. We get to NP only by allowing an unbounded number of propositions. An analogous situation occurs here (which is precisely why we allowed Φ to be infinite in general).

There is another subtlety involving how we represent formulas. Typically, when we compute upper or lower bounds on complexity for the satisfiability problem, bounds are given as functions of the length of the formula, represented as a string of symbols. Of course, there are other ways of representing the formula. We could represent it as a tree, with the leaves labeled by primitive propositions, and the interior nodes labeled by operations such as conjunction, negation, or $\Box$. With each interior node we can associate the formula that results from applying the operation labeling the node to the formulas represented by the nodes of its successors. It is easy to see that the size of the tree (i.e., the number of nodes in the tree) is proportional to the length of the original formula. Thus, choosing between these two representations is a matter of taste.

Rather than representing the formula as a tree, we could represent it as a dag (directed acyclic graph), so that a node can be the successor of more than one node. The dag representation can be exponentially more succinct than the tree representation. For example, if φ is a complex formula, then the representation of $\psi = \varphi \wedge \neg \Box \varphi$ as a dag requires only two more nodes than the representation of φ, since the node representing φ can be "reused", although the length of ψ is more than twice the length of φ. Because the dag representation is more succinct than the tree representation, a lower bound is stronger if it is proved for the tree representation, while an upper bound is stronger if it is proved for the dag representation. To distinguish the two representations, we use $||\varphi||$ to denote the length of φ under the tree representation, and $|\varphi|$ to denote the length of φ under the dag representation.

Typically not much issue is made of the representation of a formula. This is because for all logics that we are aware of, the complexity of validity is independent of whether we use a tree or dag representation. In particular, it is easy to see that this is true for propositional logic and all standard modal logics. Essentially, any upper bound for complexity that is based on considering subformulas will typically be independent of the representation. The situation is different if we consider the complexity of evaluating the truth of a propositional formula. This is known to complete for polynomial time if we take the dag representation [Lad75], and complete for alternating logarithmic time under the tree representation [Bus87].

If Φ is finite, it is easy to show that deciding almost-sure structure-validity for formulas in Φ is in polynomial time. (Of course, the constants are exponential in $|\Phi|$.) If Φ is infinite, then it is easy to show that the problem of deciding almost-sure structure validity is in $\Delta_2^p = P^{NP}$, at the second level of the polynomial hierarchy [Sto77]. This is true whether we use the tree or dag representation. Moshe Vardi has shown that a matching Δ_2^p lower bound for the dag representation. The best lower bound we have for the tree representation is hard D_2^p, the class defined in [PY82] to consist of all languages L of the form $L_1 \cap L_2$, where L_1 is in NP and L_2 is in co-NP. Note that a problem that is D_2^p-hard must also be both NP-hard and co-NP-hard. Since, as we shall see, the problem of deciding almost-sure structure validity is closely related to checking the truth of a propositional formula, it may well be that the gap here is unavoidable, although we have not been able to prove this.

Theorem 4.6: *If Φ is finite, then deciding almost-sure structure validity for formulas in $\mathcal{L}(\Phi)$ is in polynomial time (for both the dag and tree representations). If Φ is infinite, then deciding almost-sure structure validity is Δ_2^p-complete for the dag representation of formulas; it is in Δ_2^p and D_2^p hard for the tree representation.*

Proof: We can reduce the formula φ to φ^r by querying an oracle for satisfiability no more than $|\varphi|$ times. One more query will determine if φ^r is valid. These are polynomial time queries if Φ is finite, whether φ is represented as a tree or a dag; they are NP queries of Φ is infinite. Thus, it follows that almost-sure validity is in polynomial time if Φ is finite, and in Δ_2^p if Φ is infinite, for both representations. For the lower bound in the case of the tree representation, it is easy to see that the language consisting of pairs of propositional formulas (φ, ψ) such that φ is satisfiable and ψ is not (this is the language SAT-UNSAT of [PY82]) is D_2^p-complete. It follows from Proposition 4.4 that (φ, ψ) is in SAT-UNSAT iff $\neg\psi \wedge \Diamond\varphi$ is almost-surely structure valid. We leave details of the lower bound proof for the dag representation to the full paper. ∎

Finally, we can use our techniques to get a complete axiomatization for almost-sure validity. Consider the following axiom:

C. $\Diamond\varphi$ if φ is a consistent propositional formula.[1]

Let K^C be the axiom system resulting from adding axiom C to K.

Theorem 4.7: *K^C is a sound and complete axiomatization for almost-sure structure validity.*

Proof: See the full paper. ∎

As Fagin showed [Fag76], there is one (infinite) relational structure U_∞ with the property that a first-order formula (without constant or function symbols) is true in U_∞ iff it has asymptotic probability 1. An analogous property holds for structure validity:

Theorem 4.8: *Given Φ, there is a structure M_∞ such that for all formulas $\varphi \in \mathcal{L}(\Phi)$, $\nu(\varphi) = 1$ iff φ is valid in M_∞. M_∞ is finite if Φ is finite, and infinite if Φ is infinite.*

Proof: See the full paper. ∎

We have now settled the questions regarding 0-1 laws for structure validity for the modal logic K. We

[1] Of course, checking whether a given formula is an instance of this axiom scheme is NP-complete. We can get an axiom that is simpler to check (and also gives us completeness, with a little extra work) as follows. As usual, we say that a *literal* is either p or $\neg p$, where p is a primitive proposition. A *consistent conjunction of literal* is a conjunction of literals that does not contain both p and $\neg p$ as conjuncts for some primitive proposition p. Rather than considering $\Diamond\varphi$ for any consistent propositional formula φ, it is not to hard to show that it suffices to consider $\Diamond\varphi$ where φ is a consistent conjunction of literals.

can ask the same questions for the modal logics T, S4, and S5; that is, we can consider limiting probabilities with respect to $\mathcal{M}^r$, $\mathcal{M}^{rt}$, or $\mathcal{M}^{rst}$. Interestingly, it turns out that 0-1 laws coincide for all these logics. That is, a formula is almost-surely valid with respect to the structures in $\mathcal{M}^r$, $\mathcal{M}^{rt}$, or $\mathcal{M}^{rst}$ iff it is almost-surely valid with respect to the structures in $\mathcal{M}$. This follows immediately from the fact that Propositions 4.3 and 4.4 hold (with no change in proof) even if we restrict to structures in $\mathcal{M}^r$, $\mathcal{M}^{rt}$, or $\mathcal{M}^{rst}$. Thus, Theorem 4.5 holds if we replace ν_n by ν_n^r, ν_n^{rt} or ν_n^{rst}. Summarizing, we have

Theorem 4.9: *For all modal formulas φ, we have $\nu^*(\varphi) = 1$ iff the propositional formula φ^r is valid; otherwise $\nu^*(\varphi) = 0$, where ν^* is either ν^r, ν^{rt}, or ν^{rst}. Moreover, K^C is a sound and complete axiomatization for almost-sure validity with respect to $\mathcal{M}^r$, $\mathcal{M}^{rt}$, and $\mathcal{M}^{rst}$.*

5 0-1 laws for frame validity

Our main goal in this section is to prove the 0-1 law for frame validity.

Theorem 5.1: *For every modal formula φ, either $\mu(\varphi) = 0$ or $\mu(\varphi) = 1$.*

Our approach to proving Theorem 5.1 is similar to the standard tableau technique for modal satisfiability [HC78]. We define a class of frames called the *special* frames, and reduce almost-sure frame satisfiability to satisfiability in special frames. This is made precise in Theorem 5.5 below. In order to define special frames, we first need to define a few other notions.

Definition 5.2: Given a frame $F = (S, \mathcal{R})$, $s \in S$ and $A \subseteq S$, define $\mathcal{R}(s) = \{t : (s,t) \in \mathcal{R}\}$ and $\mathcal{R}(A) = \cup_{t \in A} \mathcal{R}(t)$. Similarly, define $\mathcal{R}^{-1}(s) = \{t : (t,s) \in \mathcal{R}\}$ and $\mathcal{R}^{-1}(A) = \cup_{t \in A} \mathcal{R}^{-1}(t)$. If $B \subseteq S$, we say B *$\mathcal{R}$-covers* A if $A \subseteq \mathcal{R}^{-1}(B)$. ∎

Definition 5.3: A *labeling* of a frame $(S, \mathcal{R})$ is a function f that assigns each state in S a nonnegative real number. The labeling f is *ϵ-safe* for $\epsilon \geq 0$ if, for every subset S' of states such that $\min_{s \in S'} f(s) > \epsilon$ we have

$$\sum_{s \in S'} f(s) > \sum_{(s,s') \in ((S' \times S') - \mathcal{R})} f(s) \cdot f(s'). \quad ∎$$

Definition 5.4: A frame $F = (S, \mathcal{R})$ is *ϵ-special (for $\epsilon \geq 0$) with respect to $S_0 \subseteq S$ and labeling f* if

SP1. f is ϵ-safe,

SP2. For all $T \subseteq S - S_0$, if $\sum_{t \in ((S - S_0) - T)} f(t) < 1 - \epsilon$ then $\mathcal{R}(s) \cap (S - S_0) = T$ for some $s \in S_0$,

SP3. For all $T \subseteq S - S_0$, if $\sum_{t \in T} f(t) \geq 1 + \epsilon$ then T $\mathcal{R}$-covers S_0.

A structure $M = (S, \mathcal{R}, \pi)$ is *ϵ-special for φ with respect to $S_0 \subseteq S$ and f* if (a) φ is satisfiable in M, (b) the underlying frame $(S, \mathcal{R})$ is ϵ-special with respect to S_0 and f, and (c) for all subformulas of φ of the form $\Box\psi$ and all $s \in S$, we have:

SP4. If $(M, s) \models \neg\Box\psi$, then $(M, t) \models \neg\psi$ for some $t \in S - S_0$ such $(s,t) \in \mathcal{R}$.

SP5. If $(M, s) \models \Box\psi$, then $(M, t) \models \psi$ for all $t \in S_0$.

A frame (resp., structure) is *ϵ-special* if it is ϵ-special with respect to some subset S_0 and labeling f. A frame (resp., structure) is *special (with respect to S_0 and f)* if it is 0-special (with respect to S_0 and f). Similarly, we say that a labeling is *safe* if it is 0-safe. ∎

Our interest in special structures is motivated by the following result, from which Theorem 5.1 immediately follows.

Theorem 5.5: *For any modal formula φ,*

(a) if φ is not satisfied in a finite special structure, then $\mu(\varphi) = 0$.

(b) if φ is satisfied in a finite special structure, then then $\mu(\varphi) = 1$,

Most of the rest of this section is devoted to proving Theorem 5.5. Before we get into the details of the proof, let us consider more carefully the definition of special structures. Unfortunately, we cannot provide much intuition here; the details of the definition are best motivated by the proofs we are about to present. The set S_0 can be thought of as the set where all subformulas of φ of the form $\Box\psi$ are satisfied. As we shall see, S_0 corresponds in a precise sense to a set of size $O(n)$ in almost every frame satisfying φ. Now suppose that there is a finite set of formulas $\neg\Box\psi_1, \ldots, \neg\Box\psi_k$ such that at least one of these formulas is true in every state in S_0. Let B_i be the set of states $S - S_0$ where $\neg\psi_i$ is true, $i = 1, \ldots, k$. It is easy to see that $\cup_{i=1}^k B_i$ must cover S_0. Morever, we can show that any set that covers S_0 must have size at least $\log(n) - o(\log(n))$. Thus, if $|B_i| = b_i \log(n)$, then $\sum_i b_i \geq 1$. We think of a node with label b according to the safe labeling as corresponding to a set of size $b \log(n)$, a set where a formula of the form $\neg\Box\psi$ which is true at some subset of states in S_0 is satisfied. Under this correspondence, it turns out that properties SP1–SP3 correspond to three properties that hold with probability 1 in almost all frames. We hope that further details of the definition of special structures will become clearer in the course of the proof.

Fix an integer $k > 0$ and $\delta \in (0, 1]$. Consider the following properties of a frame $F = (S, \mathcal{R})$:

F1(k, δ). For all disjoint sets $B_1, \ldots, B_l \subseteq S$ such that $l \leq k$ and $|B_i| \geq \delta \log(n)$, we have

$$\log(n) \sum_{1 \leq i \leq l} |B_i| > \sum_{\mathcal{R}(B_i) \cap B_j = \emptyset} |B_i| \cdot |B_j|.$$

F2(k, δ). For all states $u_1, \ldots, u_{k'} \in S$, and all nonempty subsets $T_1, \ldots, T_{k''}, B, C \subseteq S$ such that $k', k'' \leq k$, $T_i \cap B = \emptyset$, $1 \leq i \leq k''$, $|B| \leq (1-\delta)\log(n)$, and $|C| \geq n - k\log(n)$, we have

$$|(C \cap \mathcal{R}(u_1) \cap \ldots \cap \mathcal{R}(u_{k'}) \\ \cap \mathcal{R}^{-1}(T_1) \cap \ldots \cap \mathcal{R}^{-1}(T_{k''})) - \mathcal{R}^{-1}(B))| \geq k.$$

F3(k, δ). For all $C \subseteq S$ with $|C| \geq (1+\delta)\log(n)$, C $\mathcal{R}$-covers all of S.

Lemma 5.6: *For all $k > 0$, and $\delta \in (0, 1]$,*

$$\lim_{n \to \infty} \mu_n(F \in \mathcal{F}_n : F \textit{ satisfies F1}(k, \delta),\ k = 1, 2, 3\ ) = 1.$$

Proof: See the full paper. ∎

Lemma 5.7: *Given $\delta > 0$ and a modal formula φ, let $k = 2^{|\varphi|}$, let $F \in \mathcal{F}_n$ be a frame satisfying F1(k, δ), F2(k, δ), and F3(k, δ), such that φ is satisfiable in F. Then we can construct a structure $M^\varphi = (S^\varphi, \mathcal{R}^\varphi, \pi^\varphi)$ which is δ-special for φ with respect to $S_2 \subseteq S^\varphi$ such that $|S^\varphi| \leq 2^{2^{|\varphi|}+|\varphi|}$ and $|S^\varphi - S_2| \leq 2^{|\varphi|}$.*

Proof: Suppose that F satisfies the hypotheses of the lemma, and let $M = (S, \mathcal{R}, \pi)$ be a structure based on F in which φ is satisfiable. Roughly speaking, the idea is that we can partition S into N subsets, where $N \leq 2^{2^{|\varphi|}+|\varphi|}$. All the states in each subset agree on the truth values that they assign to subformulas of φ. Each of these subsets of states will correspond to a node in a special structure $M^\varphi = (S^\varphi, \mathcal{R}^\varphi, \pi^\varphi)$ for φ. We proceed as follows.

We define the *closure* of φ, written $cl(\varphi)$, to be the set of subformulas of φ and their negations. We say that states s and t in S are *equivalent* with respect to $cl(\varphi)$, written $s \equiv_\varphi t$, if, for every formula $\psi \in cl(\varphi)$, we have $(M, s) \models \psi$ iff $(M, t) \models \psi$. We use $[s]$ to denote the equivalence class $\{t : s \equiv_\varphi t\}$. Let S_1 be the set of equivalence classes. Note that $|S_1| \leq 2^{|\varphi|} = k$.

Suppose $|S_1| = k' \leq k$ and let $u_1, \ldots, u_{k'}$ be representatives of each equivalence class. For future reference, we call these the *canonical representatives*. Note that each equivalence class $[s] \in S_1$ contains a unique canonical representative. Define $A = \mathcal{R}(u_1) \cap \ldots \cap \mathcal{R}(u_{k'})$. From F2$(k, \delta)$ (taking $k_2 = 0$ and $B = \emptyset$), it follows that $A \neq \emptyset$. We partition A as follows: For each nonempty subset $[T]$ of S_1 and equivalence class $[s] \in S_1$, let $A_{[s],[T]}$ denote the set of all states in $[s] \cap A$ whose successors consist of precisely states in $[T]$; that is

$$A_{[s],[T]} = \{s' : s' \in [s] \cap A \cap [(\cap_{[t] \in [T]} \mathcal{R}^{-1}([t])) \\ -(\cup_{[t] \notin [T]} \mathcal{R}^{-1}([t]))]\}.$$

Define

$$S_2 = \{([s], [T]) : A_{[s],[T]} \neq \emptyset\}.$$

We take $S^\varphi = S_1 \cup S_2$. Clearly $|S^\varphi - S_2| = |S_1| \leq 2^{|\varphi|}$ and $|S_2| \leq 2^{|\varphi|}(2^{2^{|\varphi|}} - 1)$, so $|S^\varphi| \leq 2^{2^{|\varphi|}+|\varphi|}$. Let

$$\mathcal{R}^\varphi = \{([s], [t]) : [s], [t] \in S_1,\ (s, t) \in \mathcal{R}\} \\ \cup \{(([s], [T]), [u]) : [u] \in [T]\}.$$

Finally, we define $\pi^\varphi([s]) = \pi^\varphi([s], [T]) = \pi(s)$. This completes the description of M^φ.

In the full paper, we show (by induction on the structure of formulas) that if $\psi \in cl(\varphi)$ and $s \in S$, then

$(*)$ $(M, s) \models \psi$ iff $(M^\varphi, [s]) \models \psi$, and

$(**)$ if $([s], [T]) \in S_2$, then $(M, s) \models \psi$ iff $(M^\varphi, ([s], [T])) \models \psi$.

We now show that M^φ is a $(k+1)\delta$-special structure for φ with respect to S_2. The fact that φ is satisfiable in M^φ is immediate from $(*)$ and the fact that φ is satisfiable in M. For SP4, suppose that $(M^\varphi, [s]) \models \neg \Box \psi$. By $(*)$, it follows that $(M, s) \models \neg \Box \psi$. Thus, for some t such that $(s, t) \in \mathcal{R}$, we have $(M, t) \models \neg\psi$. By the definition of $\mathcal{R}^\varphi$, it follows that $([s], [t]) \in \mathcal{R}^\varphi$, and by $(*)$ again, we have that $(M^\varphi, [t]) \models \neg\psi$. Now suppose that $(M^\varphi, ([s], [T])) \models \neg \Box \psi$. By construction, there is some $s' \in [s] \cap A$ such that $(s', t') \in \mathcal{R}$ implies that $[t'] \in [T]$. Since $[s'] = [s]$, by $(**)$, we have that $(M, s') \models \neg \Box \psi$. Thus, for some t' such that $(s', t') \in \mathcal{R}$, we must have $(M, t') \models \neg\psi$. From $(*)$, we have that $(M^\varphi, [t']) \models \neg\psi$; since $[t'] \in [T]$, the definition of $\mathcal{R}^\varphi$ guarantees that $(([s], [T]), [t']) \in \mathcal{R}^\varphi$. This shows that SP4 holds.

For SP5, suppose that $(M^\varphi, [s]) \models \Box\psi$. Let u be the canonical representative in $[s]$. By construction of A, we know that $(u,t) \in \mathcal{R}$ for all $t \in A$. Thus, $(M,t) \models \psi$ for all $t \in A$. By $(**)$, it follows that for all states $([t],[T]) \in S_2$, we have $(M^\varphi, ([t],[T])) \models \psi$. If $(M^\varphi, ([s],[T]) \models \Box\psi$, by $(*)$ and $(**)$, it follows that $(M^\varphi, [s]) \models \Box\psi$, and again we get that for all states $([t],[T]) \in S_2$, we have $(M^\varphi, ([t],[T])) \models \psi$. This proves SP5.

We now must prove SP1–SP3. Define the function f on S^φ by

$$\begin{cases} f([s]) = |[s]|/\log(n) & \text{if } [s] \in S_1 \\ f(([s],[T])) = 0 & \text{if } ([s],[T]) \in S_2. \end{cases}$$

We claim that f is δ-safe. For suppose $S' \subseteq S^\varphi$ is such that $\min_{t \in S'} f(t) > \delta$. We want to show that $\sum_{t \in S'} f(t) > \sum_{(t,t') \in ((S' \times S') - \mathcal{R}^\varphi)} f(t) \cdot f(t')$. The definition of f guarantees that $S' \subseteq S_1$. Suppose $S' = \{[s_1], \ldots, [s_l]\}$. By construction, $|[s_i]| \geq \delta \log(n)$ for $i = 1, \ldots, l$. From F1(k,δ), it follows that $\log(n) \sum_{i=1}^{l} |[s_i]| > \sum_{\mathcal{R}([s_i]) \cap [s_j] = \emptyset} |[s_i]| \cdot |[s_j]|$. Note that $([s_i],[s_j]) \notin \mathcal{R}^\varphi$ iff $\mathcal{R}([s_i]) \cap [s_j] = \emptyset$. Thus, $\log(n) \sum_{i=1}^{l} |[s_i]| > \sum_{([s_i],[s_j]) \notin \mathcal{R}^\varphi} |[s_i]| \cdot |[s_j]|$. Dividing both sides of this inequality by $\log(n)^2$, and using the fact that $f([s_i]) = |[s_i]|/\log(n)$, we get the desired result. Thus, f is δ-safe, proving SP1.

For SP2, suppose that $[T] \subseteq S_1$ and

$$\sum_{[s] \in S_1 - [T]} f([s]) < 1 - \delta.$$

Let $B = \cup_{[s] \in (S_1 - [T])} [s]$. From the definition of f, we have that $|B| < (1-\delta)\log n$. From F2(k,δ), it follows that

$$(\mathcal{R}(u_1) \cap \ldots \cap \mathcal{R}(u_{k'}) \cap (\cap_{[t] \in [T]} \mathcal{R}^{-1}([t]))) - \mathcal{R}^{-1}(B) \neq \emptyset,$$

where $u_1, \ldots, u_{k'}$ are the canonical representatives, as chosen above. But this means that for some $s \in A$, we have $A_{[s],[T]} \neq \emptyset$. Thus, $([s],[T]) \in S_3$. By definition, $\mathcal{R}^\varphi(([s],[T]) = [T]$. This proves SP2.

For SP3, suppose that $[T] \subseteq S_1$ and $\sum_{[s] \in [T]} f([s]) \geq 1 + \delta$. Let $C = \cup_{[t] \in [T]} [t]$. From the definition of f we have that $|C| \geq (1+\delta)\log(n)$. Thus, from F3$(k,\delta)$, we get that C $\mathcal{R}$-covers S. It easily follows that $[T]$ $\mathcal{R}^\varphi$-covers S_2. This proves SP3.

Thus, we have shown that M^φ is a δ-special structure for φ. ∎

We have just shown that under appropriate assumptions, there exists a δ-special structure for φ. Now we want to strengthen this to get a special (i.e., a 0-special) structure for φ. We first need a technical result about safe labelings.

Lemma 5.8: *If F is ϵ-special with respect to S_0 and f, then F is also ϵ-special with respect to S and f', where $f'(s) = \min(f(s), 1)$ for $s \in S - S_0$ and $f'(s) = 0$ for $s \in S_0$.*

Part (a) of Theorem 5.5 now follows immediately from the following result.

Proposition 5.9: *If $\mu(\varphi) \neq 0$, then there is a special structure $M = (S, \mathcal{R})$ for φ with respect to a set $S_0 \subseteq S$ such $|S| \leq 2^{2^{|\varphi|} + |\varphi|}$ and $|S - S_0| \leq 2^{|\varphi|}$.*

Proof: Suppose $\mu(\varphi) \neq 0$. Then there exist infinitely many n such that

$$\mu_n(\{F \in \mathcal{F}_n : \text{ some structure based on } F \text{ satisfies } \varphi\})$$

is strictly greater than 0. From Lemma 5.6 and the hypotheses of the proposition, it follows that there exists an increasing sequence $n_1, n_2, \ldots$ such that for all m, there exists a frame with n_m states such that the hypotheses of Lemma 5.7 hold with $\delta = 1/m$. Thus, corresponding to each n_m there exists a $(1/m)$-special structure $M_m = (S_m, \mathcal{R}_m, \pi_m)$ for φ with respect to a subset $S'_m \subseteq S_m$ and a safe labeling f_m such that $|S_m| \leq 2^{2^{|\varphi|}+|\varphi|}$ and $|S'_m - S_m| \leq 2^{|\varphi|}$. Since there are only finitely many structures satisfying φ with at most $2^{2^{|\varphi|}+|\varphi|}$ states, there must exist one structure $M = (S, \mathcal{R}, \pi)$ and a subset $S_0 \subseteq S$ such that $M = M_m$ and $S_0 = S'_m$ for infinitely many m. Thus, M is a $(1/m)$-special structure for φ with respect to S_0 and f_m for all m. We want to show that M is in fact special.

By Lemma 5.8, we can assume without loss of generality that $f_m(s) \leq 1$ for all $s \in S$. Suppose that $S = \{s_1, \ldots, s_N\}$. The sequence of tuples $(f_0(s_1), \ldots, f_0(s_N)), (f_1(s_1), \ldots, f_1(s_N)), \ldots$ is a sequence in the compact space $[0,1]^N$. Thus, the sequence has an accumulation point $(a_1, \ldots, a_N)$; i.e., a subsequence of this sequence converges to $(a_1, \ldots, a_N)$. Now consider the labeling defined by $f(s_i) = a_i$. Straightforward continuity arguments show that M is a special structure of φ with respect to f. We leave details to the reader. ∎

In order to prove Part (b) of Theorem 5.5, we show that a special structure can be embedded in almost every frame in such a way as to preserve satisfiability.

Suppose that $F = (S, \mathcal{R})$ and $F' = (S', \mathcal{R}')$ are two frames such that $|S| \leq |S'|$, and suppose $S_0 \subseteq S$. We say that F is *S_0-embeddable in F'* if there is an onto mapping $\gamma : S' \rightarrow S$ such that

P1. if $(s,t) \in \mathcal{R}$ and $t \notin S_0$, then $\gamma^{-1}(t)$ $\mathcal{R}$-covers $\gamma^{-1}(s)$, P2. if $(s,t) \notin \mathcal{R}$ and $t \notin S_0$, then $\mathcal{R}'(\gamma^{-1}(s)) \cap \gamma^{-1}(t) = \emptyset$.

Lemma 5.10: *If $F = (S, \mathcal{R})$, $S_0 \subseteq S$, M is a special structure for φ with respect to S_0 based on F, and F is S_0-embeddable in F', then φ is satisfied in F'.*

Proof: See the full paper. ∎

Part (b) of Theorem 5.5 now follows immediately from the following result.

Proposition 5.11: *If F is a finite special frame with respect to S_0, then*

$$\lim_{n\to\infty} \mu_n(F' \in \mathcal{F}_n : F \textit{ is } S_0\textit{-embeddable in } F') = 1.$$

Proof: Suppose $F = (S, \mathcal{R})$ is special with respect to $S_0 \subseteq S$ and f. By Lemma 5.8, without loss of generality, we can assume that $f(s) \leq 1$ for all $s \in S$. Let $S_1 = S - S_0$. Our first step is to show that for almost all frames $F' = (S', \mathcal{R}')$, we can find an onto function $\gamma : S' \to S$ such that

P1$'$. if $(s,t) \in ((S_1 \times S_1) \cap \mathcal{R})$, then $\gamma^{-1}(t)$ $\mathcal{R}$-covers $\gamma^{-1}(s)$,

P2$'$. if $(s,t) \in ((S_1 \times S_1) - \mathcal{R})$, then $\mathcal{R}'(\gamma^{-1}(s)) \cap \gamma^{-1}(t) = \emptyset$.

P3. if $s \in S_1$ and $f(s) > 0$, then $|\gamma^{-1}(s)| = \lfloor f(s)\log(n) \rfloor$

P4. if $f(s) = 0$ then $|\gamma^{-1}(s)| = 1$.

P5. if $T \subseteq S_1$ and $\sum_{t\in T} f(t) \geq 1$, then $\gamma^{-1}(T)$ $\mathcal{R}'$-covers $\gamma^{-1}(S_0)$.

Notice that P1$'$ and P2$'$ are weaker variants of P1 and P2, respectively, where we only focus on pairs (s,t) in $S_1 \times S_1$.

In order to prove this, we use the *second moment method*, a standard technique in random graph theory [Bol85]. Let X be a random variable on $\mathcal{F}$ such that $X(F')$ is the number of mappings γ from F' to F satisfying P1$'$, P2$'$, P3, P4, and P5. We want to show

(†) $\lim_{n\to\infty} \mu_n(\{F' \in \mathcal{F}_n : X(F') > 0\}) = 1$.

Let $E_n(X)$ be the expected value of $X(F')$ in $\mathcal{F}_n$. According to the second moment method, to prove (†), it suffices to show that $\lim_{n\to\infty} E_n(X) = \infty$ and $\lim_{n\to\infty} E_n(X^2)/E_n(X)^2 = 1$. We do this in the full paper.

Once we have found γ which satisfies, P1$'$, P2$'$, P3, P4 and P5, we are almost done. Choose ϵ so that for all $T \subseteq S_1$, if $\sum_{t\in T} f(t) < 1$, then in fact $\sum_{t\in T} f(t) < 1 - \epsilon$. Let $N = |S_0| 2^{|S_1|}$. By Lemma 5.6, the asymptotic probability that a frame satisfies F2(N,ϵ) is 1. Consider a frame $F' = (S', \mathcal{R}')$ satisfying F2(N,ϵ) such that there is a mapping $\gamma : F' \to F$ satisfying P1$'$, P2$'$, P3, P4, and P5. Let $S'_0 = \gamma^{-1}(S_0)$, and let $S'_1 = \gamma^{-1}(S_1)$. Notice that $S'_0 = S' = S'_1$. We now show that, in fact, by possibly redefining γ on S'_0, we can get a map $\gamma' : F' \to F$ satisfying P1 and P2. This shows that F is embeddable in F'.

For each $T \subseteq S_1$, let S_T consist of all the states in S_0 with $\mathcal{R}$-successors in T and no $\mathcal{R}$ successors in $S_1 - T$; similarly, let S'_T consist of all the states in S'_0 with $\mathcal{R}'$-successors in $\gamma^{-1}(t)$ for $t \in T$, and no $\mathcal{R}'$ successors in $\gamma^{-1}(t)$ for $t \in S_1 - T$. That is,

$$S_T = (S_0 \cap (\cap_{t\in T} \mathcal{R}^{-1}(t))) - \cup_{t\in S_1 - T} \mathcal{R}^{-1}(t)$$

and

$$S'_T = (S'_0 \cap (\cap_{t\in T} \mathcal{R}'^{-1}(\gamma^{-1}(t)))) - \cup_{t\in S_1 - T} \mathcal{R}'^{-1}(\gamma^{-1}(t)).$$

If $\sum_{t\in S_1 - T} f(t) < 1$, then by SP2, $S_T \neq \emptyset$ and by F2(N,ϵ) and the choice of N, we have $|S'_T| \geq N \geq |S_0|$. If $\sum_{t\in S_1 - T} f(t) \geq 1$, then by SP3, $S_T = \emptyset$, and by P5, $S'_T = \emptyset$.

It is straightforward to redefine γ on S'_0 to get an onto map $\gamma' : F' \to F$ that agrees with γ on S'_1 such that $\gamma'^{-1}(S_T) = S'_T$. We now show that γ' satisfies P1 and P2. We start with P1. If $(s,t) \in \mathcal{R}$ and both $s,t \in S_1$, since γ' agrees with γ on S'_1, it follows from P1$'$ that $\gamma'^{-1}(t)$ covers $\gamma'^{-1}(s)$. Suppose that $s \in S_0$. Then $s \in S_T$ for some $T \subseteq S_1$ such that $t \in T$. By construction, $\gamma'^{-1}(s) \in S'_T$ and $\gamma'^{-1}(t)$ covers $\gamma' - 1(s)$. This shows that P1 holds.

For P2, suppose that $(s,t) \notin \mathcal{R}$. If both $s,t \in S_1$, then it follows from P2$'$ that $\mathcal{R}(\gamma'^{-1}(s)) \cap \mathcal{R}(\gamma' - 1(t)) = \emptyset$. If $s \in S_0$, then $s \in S_T$ for some $T \subseteq S_1$ such that $t \notin T$. By construction, $\gamma'^{-1}(s) \in S'_T$, and so $\mathcal{R}'(\gamma'^{-1}(s)) \cap \gamma'^{-1}(t) = \emptyset$. This shows that P2 holds.

Thus, we can embed F into a frame that satisfies F2(N,ϵ) for which there is a mapping γ satisfying P1$'$, P2$'$, P3, P4, and P5. Since these are both properties that hold with asymptotic probability 1, it follows that $\lim_{n\to\infty} \mu_n(\{F' \in \mathcal{F}_n : F \text{ is } S_0\text{-embeddable in } F'\}) = 1$. ∎

We have now completed the proof of Theorem 5.5. The theorem and its proof gives us a great deal of information, which we can now exploit. For one thing, we get an analogue of Theorem 4.8 for frames:

Theorem 5.12: *Given Φ, there is an infinite frame M_∞ such that for all formulas $\varphi \in \mathcal{L}(\Phi)$, $\mu(\varphi) = 1$ iff*

φ is valid in F_∞.

Proof: Just take F_∞ to be the disjoint union of all special frames. ▮

Note that F_∞ is infinite, even if Φ is finite.

Turning our attention to complexity, we can show

Theorem 5.13: *For both the dag and the tree representations, and both finite and infinite sets of propositions Φ, deciding almost-sure frame-validity of formulas in $\mathcal{L}(\Phi)$ is is double-exponential space, and is exponential-time hard.*

Proof: The lower bound follows much the same lines as the exponential time lower bound of Fischer and Ladner [FL79] for one-letter PDL. We leave details to the full paper.

For the upper bound, the problem reduces to deciding whether there is a special structure for φ. We know that if there is one at all, then there is a special structure $M = (S, \mathcal{R}, \pi)$ for φ with respect to S_0 such that $|S| \leq 2^{2^{|\varphi|}+|\varphi|}$ and $|S - S_0| \leq 2^{|\varphi|}$. For each structure $M = (S, \mathcal{R}, \pi)$ and subset $S_0 \subseteq S$ satisfying these constraints, we can easily check if (in double-exponential time) if φ is satisfiable in M and properties SP4 and SP5 hold. We now need to check that properties SP1–SP3. This amounts to checking whether there is a safe labeling satisfying SP2 and SP3. Without loss of generality (using Lemma 5.8), we can assume that the safe labeling assigns value 0 to all the states in S_0. Thus, all that matters is the value assigned to the (at most exponential number of) states in $S - S_0$. We can find a formula in the language of arithmetic (i.e., over $+, \times, 0, 1$), whose length is at most double-exponential in $|\varphi|$, that is valid in the theory of real closed fields iff such a safe mapping exists: If $S - S_0 = \{s_1, \ldots, s_N\}$ (where $N \leq 2^{|\varphi|}$), then we simply say there exist numbers $x_1, \ldots, x_N$ such that the appropriate properties hold, where the "appropriate properties" can easily be expressed by a double-exponential length quantifier-free formula in $x_1, \ldots, x_N$ (that depends on the $\mathcal{R}$ relation in M). Since this formula has only existential quantifiers, by a result of Canny [Can88], we can check whether it is true in double-exponential space. Thus, in double-exponential space, we can check whether a special structure for φ exists. ▮

Just as for structure validity, we can also ask whether a 0-1 law holds for frame validity if we consider frames in $\mathcal{F}^r$, $\mathcal{F}^{rt}$, or $\mathcal{F}^{rst}$. The answer in all cases is yes. While this is may not seem surprising in light of the results we have already proved, what is perhaps a surprise is that, in the case of $\mathcal{F}^{rt}$ and $\mathcal{F}^{rst}$, almost-sure frame validity is equivalent to S5-validity.

Theorem 5.14: *For every modal formula φ*

(a) $\mu^r(\varphi) = 1$ if φ if there is a special modal structure $M = (S, \mathcal{R}, \pi)$ for φ such that $\mathcal{R}$ is reflexive; otherwise $\mu^r(\varphi) = 0$.

(b) $\mu^{rt}(\varphi) = \mu^{rst}(\varphi) = 1$ if φ is S5-valid; otherwise $\mu^{rt}(\varphi) = \mu^{rst}(\varphi) = 0$.

Proof: The proof for part (a) follows similar lines to that of the proof of Theorem 5.5; we omit details here. For part (b), we use the observation that with asymptotic probability 1, all states are connected by a path of length 2. Thus, if $\mathcal{R}$ is transitive, then $\mathcal{R}$ is almost-surely the universal relation. It follows that all S5-valid formulas are almost-surely valid in $\mathcal{F}^{rt}$ and $\mathcal{F}^{rst}$. For the converse, suppose that φ is S5-satisfiable. Then there is a finite structure $M = (S, \mathcal{R}, \pi)$ satisfying φ such that $\mathcal{R}$ is universal. It is easy to see that if $M' = (S', \mathcal{R}', \pi')$ is such that $\mathcal{R}'$ is universal and $\{\pi'(s') : s' \in S'\} = \{\pi(s) : s \in S\}$, then φ is satisfiable in S'. Since we have already observed that, in almost all structures in $\mathcal{F}^{rt}$ and $\mathcal{F}^{rst}$, the relation is universal, it easily follows that φ is satisfiable in almost all frames in $\mathcal{F}^{rt}$ and $\mathcal{F}^{rst}$. ▮

Acknowledgements: We would like to thank Moshe Vardi for his many useful comments on the paper, and for allowing us to include his result on Δ_p^2-hardness (Theorem 4.6) and his proof that frame satisfiability can express SAT (Theorem 3.3). Ron Fagin also provided useful comments on a draft of the paper.

References

[AU79] A. V. Aho and J. D. Ullman. Universality of data retrieval languages. In *Proc. 6th ACM Symp. on Principles of Programming Languages*, pages 110–117, 1979.

[Ben83] J. F. A. K. van Benthem. *Modal Logic and Classical Logic.* Bibliopolis, 1983.

[Bol85] B. Bollobas. *Random Graphs.* Academic Press, 1985.

[Bus87] S. R. Buss. The Boolean formula value problem is in ALOGTIME. In *Proc. 19th ACM Symp. on Theory of Computing*, pages 123–131, 1987.

[Can88] J. F. Canny. Some algebraic and geometric computations in PSPACE. In *Proc. 20th ACM Symp. on Theory of Computing*, pages 460–467, 1988.

[Com88] K. Compton. 0-1 laws in logic and combinatorics. In I. Rival, editor, *Proc. 1987 NATO Adv. Study Inst. on algorithms and order*, pages 353–383, Dordrecht, Netherlands, 1988. Reidel.

[DG79] B. Dreben and W. D. Goldfarb. *The Decision Problem: Solvable Classes of Quantificational Formulas.* Addison-Wesley, Reading, MA, 1979.

[Fag74] R. Fagin. Generalized first-order spectra and polynomial-time recognizable sets. In R. M. Karp, editor, *Complexity of Computation, SIAM-AMS Proceedings, Vol. 7*, pages 43–73, 1974.

[Fag75] R. Fagin. Monadic generalized spectra. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 21:89–96, 1975.

[Fag76] R. Fagin. Probabilities on finite models. *Journal of Symbolic Logic*, 41(1):50–58, 1976.

[FL79] M. J. Fischer and R. E. Ladner. Propositional dynamic logic of regular programs. *Journal of Computer and System Sciences*, 18(2):194–211, 1979.

[GHK92] A. J. Grove, J. Y. Halpern, and D. Koller. Asymptotic conditional probabilities for first-order logic. In *Proc. 24th ACM Symp. on Theory of Computing*, 1992.

[GKLT69] Y. V. Glebskiĭ, D. I. Kogan, M. I. Liogon'kiĭ, and V. A. Talanov. Range and degree of realizability of formulas in the restricted predicate calculus. *Kibernetika*, 2:17–28, 1969.

[Gra83] E. Grandjean. Complexity of the first–order theory of almost all structures. *Information and Control*, 52:180–204, 1983.

[Hal87] J. Y. Halpern. Using reasoning about knowledge to analyze distributed systems. In J. F. Traub, B. J. Grosz, B. W. Lampson, and N. J. Nilsson, editors, *Annual Review of Computer Science, Vol. 2*, pages 37–68. Annual Reviews Inc., Palo Alto, CA, 1987.

[HC78] G. E. Hughes and M. J. Cresswell. *An Introduction to Modal Logic.* Methuen, London, 1978.

[HC84] G. E. Hughes and M. J. Cresswell. *A Companion to Modal Logic.* Methuen, London, 1984.

[Kau87] M. Kaufmann. Counterexample to the 0-1 law for existential monadic second-order logic. CLI Internal Note 32, Computational Logic Inc., December 1987.

[KV87] Ph. G. Kolaitis and M. Y. Vardi. The decision problem for the probabilities of higher-order properties. In *Proc. 19th ACM Symp. on Theory of Computing*, pages 425–435, 1987.

[KV90] Ph. G. Kolaitis and M. Y. Vardi. 0-1 laws and decision problems for fragments of second-order logic. *Information and Computation*, 87:302–338, 1990.

[Lad75] R. E. Ladner. The circuit value problem is log-space complete for p. *SIGACT News*, 6(2):18–20, 1975.

[Lad77] R. E. Ladner. The computational complexity of provability in systems of modal propositional logic. *SIAM Journal on Computing*, 6(3):467–480, 1977.

[MP81] Z. Manna and A. Pnueli. Verification of temporal programs: the temporal framework. In R. S. Boyer and J. S. Moore, editors, *The Correctness Problem in Computer Science.* Academic Press, New York, 1981.

[PS89] L. Pacholski and W. Szwast. The 0-1 law fails for the class of existential second-order Gödel sentences with equality. In *Proc. 30th IEEE Symp. on Foundations of Computer Science*, pages 280–285, 1989.

[PY82] C. H. Papadimitriou and M. Yannakakis. The complexity of facets (and some facets of complexity). *Journal of Computer and System Sciences*, 28(2):244–259, 1982.

[Sto77] L. J. Stockmeyer. The polynomial-time hierarchy. *Theoretical Computer Science*, 3:1–22, 1977.