Arch. Math. Logic 42, 221–243 (2003)
Digital Object Identifier (DOI):
10.1007/s001530100135

Mathematical Logic

Valentin Goranko · Bruce Kapron

The modal logic of the countable random frame

Received: 1 May 2000 / Revised version: 29 July 2001 /
Published online: 2 September 2002 – © Springer-Verlag 2002

Abstract. We study the modal logic ML^r of the countable random frame, which is contained in and 'approximates' the modal logic of almost sure frame validity, i.e. the logic of those modal principles which are valid with asymptotic probability 1 in a randomly chosen finite frame. We give a sound and complete axiomatization of ML^r and show that it is not finitely axiomatizable. Then we describe the finite frames of that logic and show that it has the finite frame property and its satisfiability problem is in EXPTIME. All these results easily extend to temporal and other multi-modal logics. Finally, we show that there are modal formulas which are almost surely valid in the finite, yet fail in the countable random frame, and hence do not follow from the extension axioms. Therefore the analog of Fagin's transfer theorem for almost sure validity in first-order logic fails for modal logic.

1. Introduction

This paper is intended for readers mainly familiar with the modal logic side of the story, and not expected to have background on random structures and asymptotic probabilities, so we offer a brief introduction to the subject, emphasizing on 0-1 laws and almost sure validity in finite structures.

The studies of random structures apparently began from the work of Erdös and Renyi on random graphs and applications of probabilistic methods in combinatorics. The logical trend in these studies goes back (at least) to Carnap who proposed in [Carnap 50] the idea of considering asymptotic probabilities of properties expressible by logical formulas. Furthermore, Carnap proved the following remarkable result for the first-order language L_1 containing only unary predicate symbols: *the proportion of L_1-structures (taken up to isomorphism) with n elements in which a given L_1-sentence is true tends either to 0 or to 1 as n tends to infinity.* In probabilistic terms, this is the so called *(unlabelled) zero-one law* for L_1: every first-order definable property of unary relational structures holds in a randomly

V. Goranko: Department of Mathematics, Rand Afrikaans University, PO Box 524, Auckland Park 2006, Johannesburg, South Africa

B. Kapron: Department of Computer Science, University of Victoria, British Columbia, Canada

Mathematics Subject Classification (2000): 03B45, 03B70, 03C99

Key words or phrases: Modal logic – Random frames – Almost sure frame validity – Countable random frame – Axiomatization – Completeness

chosen finite structure with probability either 0 or 1, i.e. it is either *almost surely valid* or *almost surely invalid* in the finite.

Later, Gaifman studied in [Gaifman 64] infinite random structures as probabilistic models for arbitrary relational first-order languages and proved that the first-order theory of such structures satisfies an infinite set of *extension axioms*: sentences claiming that every n-tuple in the structure can be extended to an $(n + 1)$-tuple in all possible (i.e. consistent) ways. Furthermore, he showed that the first-order theory EXT axiomatized with all extension axioms is ω-categorical and complete. The probabilistic aspect of this result is rather surprising: it means that, assuming uniform distribution, any randomly constructed countable relational structure is isomorphic with probability 1 to a certain structure, *the countable random structure*! For graphs this fact seemingly had already been noted by Radó and Erdös, but Gaifman's result showed that for any relational language this ubiquitous structure is precisely the unique, up to isomorphism, countable model of EXT.

The 0-1 law for arbitrary relational first-order languages was first proved in [Glebskii et al 69]. It was established by an "almost sure" quantifier elimination proved by an involved induction on the complexity of the formula but was published in a not widely available journal. The result, and the paper itself, remained largely unnoticed until a few years later Fagin, unaware of that work, but led by attempts to solve the spectrum problem, noticed the intimate connection between Carnap's and Gaifman's results and rediscovered the 0-1 law for first-order logic, to which he gave an elegant and insightful proof in [Fagin 76]. He noticed that every finite set of extension axioms not only has a finite model, but is satisfied by *almost every* finite structure! Thus, Fagin provided two purely logical descriptions of the almost surely valid first-order properties, viz. he showed that the following are equivalent:

i) ϕ is almost surely true in finite L-structures.
ii) ϕ follows from (finitely many) extension axioms.
iii) ϕ is true in the countable random structure.

These results sparked extensive and fruitful research on asymptotic probabilities of properties formalizable in logical languages. The 0-1 law was proved: for the extension of first-order logic with a transitive closure operator in [Talanov 81]; for the more expressive extension of first-order logic with a fixed point operator in [Blass, Gurevich and Kozen 85]; later subsumed by the 0-1 law for the infinitary logic over bounded number of variables $L^{\omega}_{\infty,\omega}$ proved in [Kolaitis and Vardi 92]; for some prefix-defined fragments of monadic second-order logic in [Kolaitis and Vardi 90] who also established strong relations between decidability and 0-1 laws of such fragments; for modal logic in [Halpern and Kapron 94].

In general, however, the 0-1 law turns out to be rather a rare phenomenon than a rule. It can be easily seen that the presence of a single constant in the language is fatal for it; still, it was proved in [Lynch 85] that every sentence in a first-order language with only unary functions does have an *asymptotic probability* (though in general not 0 or 1). For second-order logic the 0-1 law fails badly, as even its monadic existential fragment contains sentences with no asymptotic probability, as first proved by Kaufmann (see [le Bars 99a] for a very accessible account of Kaufmann's counterexample, and [le Bars 99b], and [Kolaitis and Vardi 89] for more). While for prefix-defined fragments of monadic Σ^1_1 the boundary of 0- 1 laws seems to be already essentially delineated (see [Kolaitis and Vardi 89] for a survey, and [le Bars 99b] for recent additions), in general it seems rather less known, especially for the full (monadic) second-order logic.

For more on 0-1 laws and asymptotic probabilities, besides the references above, see the classical survey [Compton 88], and the very readable [Gurevich 92].

An important problem in the area, relatively independent from the truth or otherwise of 0-1 laws, is to give a logical characterization of the almost sure truth/validity in the finite for the various logical languages. The result of Fagin, mentioned above, is the definitive solution to that problem for first-order logic. Interestingly, the *0-1 law-via-transfer* phenomenon, viz. the property of the countable random structure to satisfy precisely those sentences which are almost surely valid in the finite, persists in the infinitary logic, as well as in all so far known prefix-defined fragments of second-order logic for which the 0-1 law holds. Therefore, in all these cases, the almost surely valid properties are those which follow from the extension axioms. This, however, is not always the case, even for Σ^1_1: for instance it was shown in [Blass, Harary 79] that the property of a graph to be Hamiltonian (which is NP-complete, hence Σ^1_1-definable) is almost surely valid but does not follow from any extension axiom. Furthermore, as results in [Dawar and Grädel 95] indicate, there is no natural, in sense of abstract model theory, logic which can express Hamiltonicity and has the 0-1 law.

However, even in cases when there is no transfer theorem, or the 0-1 law fails hopelessly, the class of sentences almost surely true in the finite is coherently determined and it is natural to pursue its logical characterization, be it model-theoretic or axiomatic. Besides the obvious practical importance of this problem (especially for logics with high complexity, or undecidable), it turns out that the transition from 'absolute truth' to 'truth with probability 1' can reduce the complexity qualitatively, witness first-order logic where almost sure validity in the finite, as proved in [Grandjean 83], is PSPACE-complete, vs. the undecidable validity in the finite, by Trachtenbrot's theorem.

Apart from the cases of 0-1 law via transfer, there seem to be very few results describing logically almost sure validity in the finite. An obvious reason is that this is not an *a priori* logically tractable concept. It lacks explicit logical semantics in terms of single models, but rather involves the class of all finite models as a whole, so it is *essentially global* unlike absolute validity. It is conceivable, for instance, that almost sure validity can be of such a high complexity that a recursive axiomatization is impossible, while there are no finite models of all almost sure validities. On the other hand, it is semantically well-behaved as it contains all valid formulas and is closed under finitary logical consequence, and it is deductively well behaved, being closed under finitary rules of inference, such as modus ponens: if A and $A \to B$ are true in almost every finite model, then so is B.

In modal logic there are two basic notions of validity: in models and in frames. Respectively, there are two concepts of almost sure validity. [Halpern and Kapron 94] prove the 0-1 law for both of them. The result for *almost sure model validity* (called in [Halpern and Kapron 94] “structure validity") follows easily from Fagin’s theorem, as model validity can be expressed in first-order logic by means of van Benthem’s standard translation (see [van Benthem 85]), but [Halpern and Kapron 94] does more: it provides an easy algorithm for checking if a modal formula is almost surely structure valid, and gives an explicit complete axiomatization of that concept, which turns out to be captured precisely by Carnap’s logic, the axioms of which are all $\Diamond\phi$ where ϕ is a consistent propositional formula. Note that this is not a normal modal logic, as it is not closed under substitutions. Regarding *almost sure frame validity*, however, the result is essentially new and the proof is a rather complicated combinatorial estimation of probabilities, which unfortunately offers little logical insight.[1] Furthermore, the proof implies a kind of a transfer theorem which associates almost sure frame validity with validity in a certain infinite frame which is a disjoint union of *special finite frames* of rather inexplicit nature, so this transfer result does not provide much insight into the nature of the modal logic of almost sure frame validity either. On the other hand, results in the present paper show that there is no transfer theorem with respect to the *countable random frame* F^r, and strongly indicate that probably there is no “natural" transfer theorem at all.

On the other hand, an easy compactness argument due to Kolaitis and Vardi shows that every Π^1_1 sentence which is valid in F^r follows from an extension axiom and hence is almost surely valid in finite frames. Therefore, the normal modal logic $\mathbf{ML}^r$ of F^r is contained in the normal modal logic $\mathbf{ML}^{as}$ of almost surely frame-valid formulas. As we show here, the inclusion is proper.

In this paper we study the logic $\mathbf{ML}^r$. Unlike the intractable logic of almost sure frame-validity, $\mathbf{ML}^r$ being determined by a single, though infinite, frame has a perspicuous semantics and turns out to be a reasonably well-behaved, though non-finitely axiomatizable, canonical modal logic with the finite frame property. The major difficulty in the completeness proof is to show that the canonical model satisfies the same formulas as the infinite random frame. The proof, however, is quite portable and easily extends to temporal and various other extended multi-modal logics.

Thus, the paper can be regarded as a study of the *well-behaved fragment* of almost sure frame validity in modal logic, if not the whole of it.

The paper is organized as follows: The preliminary section 1 introduces formally the basic concepts and facts related to random frames and almost sure validity and proves some results in modal logic needed later. Section 2 gives a more detailed background on the logic $\mathbf{ML}^r$ and contains some results on almost sure validity. In section 3 we give a canonical axiomatization of $\mathbf{ML}^r$ and prove its completeness with respect to F^r. We also prove that $\mathbf{ML}^r$ is not finitely axiomatizable, but has the finite frame property via filtration, and the finite models are easily recognizable, which establishes an EXPSPACE upper bound for the complexity of its satisfiability. Finally, we discuss special modally definable properties expressing non-existence of certain partitions, called *kernel partitions* in random frames,

[1] In a recent paper [le Bars 02] Le Bars claims that this result is wrong, by providing a counter-example. Anyway, none of the results in our paper depends on the truth or otherwise of the 0-1 law for frame validity of modal formulae.

and show that there are modal formulas corresponding to simple such partitions, which are almost surely valid in the finite, yet fail in the countable random frame, and hence do not follow from the extension axioms. The paper ends with some concluding remarks.

2. Preliminaries

2.1. Modal logic

We assume basic familiarity with the syntax and Kripke semantics of standard normal modal logics, including the notions of *canonical model, p-morphism, bisimulation*. For information on these, see e.g. [van Benthem 85] or [Blackburn, de Rijke, and Venema 2001].

We will need the following notion of **bounded bisimulation**, introduced first by Hennessy and Milner.

Definition 1. *Let* $M_1 = \langle W_1, R_1, V_1\rangle$ *and* $M_2 = \langle W_2, R_2, V_2\rangle$ *be models, and* $x_1 \in W_1, x_2 \in W_2$. *We define the property of* (M_1, x_1) *and* (M_2, x_2) *to be* k-**bisimilar**, *denoted* $(M_1, x_1) \sim_k (M_2, x_2)$ *inductively on* $k \in N$ *as follows:*

- $(M_1, x_1) \sim_0 (M_2, x_2)$ *iff* x_1 *and* x_2 *satisfy the same propositional variables.*
- $(M_1, x_1) \sim_{k+1} (M_2, x_2)$ *if* $(M_1, x_1) \sim_k (M_2, x_2)$ *and:*
  (forth)*: For every* $y_1 \in M_1$ *such that* $R_1x_1y_1$ *there exists a* $y_2 \in M_2$ *such that* $R_2x_2y_2$ *and* $(M_1, y_1) \sim_k (M_2, y_2)$*;*
  (back)*: Similarly for* M_1 *and* M_2 *exchanged.*

In terms of games, k-bisimulation corresponds to existence of a winning strategy for the duplicator in a k-round 2-pebble game.

k-bisimulation between rooted models can be described logically by the following modal analogs of Fraïssé formulas in first-order logic.

Definition 2. *Let* L *be a modal language with a finite set of propositional variables* $P = \{p_1, ..., p_n\}$, $M = \langle W, R, V\rangle$ *be an* L*-model, and* $x \in W$. *For every* $k \in N$ *we define inductively on* k *the* **modal k-type of** x **in** M, $T_k(x, M)$ *as follows.*

- $T_0(x, M) = \bigwedge\{p | x \in V(p)\} \wedge \bigwedge\{\neg p | x \notin V(p)\}$.
  Note that there are finitely many 0-types in L.
- $T_{k+1}(x, M) = T_k(x, M) \wedge \bigwedge\{\Diamond T_k(y, M) | Rxy\} \wedge \Box \bigvee\{T_k(y, M) | Rxy\}$.
  Assuming there are only finitely many k-types in L it follows that the conjunctions and disjunctions in this definition are finite, and hence there are only finitely many $(k + 1)$-types, too.

Proposition 1. *Let* L *be a modal language with a finite set of variables,* $M_1 = \langle F_1, R_1, V_1\rangle$ *and* $M_2 = \langle F_2, R_2, V_2\rangle$ *be* L*-models, and* $x_1 \in F_1, x_2 \in F_2$. *Then for any* $k \in N$ *the following are equivalent:*

1. $(M_1, x_1) \sim_k (M_2, x_2)$.
2. $T_k(x_1, M_1) \equiv T_k(x_2, M_2)$.
3. (M_1, x_1) *and* (M_2, x_2) *satisfy the same formulas of modal depth at most* k.

Proof.

(1) $\Leftrightarrow$ (2): Straightforward induction on k.

(2) $\Leftrightarrow$ (3): Again, induction on k. Note that $T_k(x, M)$ has modal depth k and that every modal formula of depth $k + 1$ is equivalent to a disjunction of formulae $\psi \wedge \Box\phi_0 \wedge \bigwedge_{i=1}^{m} \Diamond\phi_i$ where the modal depth of ψ is 0 and that of each of $\phi_0, \phi_1, ..., \phi_m$ is at most k. $\Box$

2.2. Random Kripke frames

Given $n \in N$, a **random (labelled) frame of size n** is a frame obtained by random and independent assignments of truth/falsity of the binary relation on every pair (x, y) from the set $\{1, \dots, n\}$, with probability for truth $p(n)$. In this paper we consider $p(n)$ to be a constant. In the particular case when $p = 0.5$, the random frame can be obtained by a random assignment of a binary relation on the domain, using uniform distribution, i.e. considering all possible binary relations equiprobable. However, the results used and those obtained here hold for any constant probability $p \in (0, 1)$.

The probability space on all n-element frames constructed as above will be denoted by $\mathcal{S}(n, p)$.

In this sub-section we consider frames as first-order structures for a language with = and one binary relational symbol.

For any property of frames $\mathcal{P}$ we denote by $\mu_{n,p}(P)$ the probability of P in $\mathcal{S}(n, p)$, i.e. the probability that P holds for a randomly constructed n-element frame. In particular, if ϕ is a first-order sentence, $\mu_{n,p}(\phi)$ will denote the probability for ϕ to be true in a frame from $\mathcal{S}(n, p)$. Note that these are discrete probabilities since $\mathcal{S}(n, p)$ is finite.

Now, we consider $\mu_p(\phi) = \lim_{n\to\infty} \mu_{n,p}(\phi)$ and if it exists we call it **the asymptotic probability** of ϕ. Once the probability p is fixed, we omit the subscript.

A sentence ϕ is said to be *almost surely valid in the finite* if $\mu(\phi) = 1$, and respectively, *almost surely invalid* if $\mu(\phi) = 0$.

Remark 1.

1) Note that the probability measure μ defined as above is not countably additive. E.g, $\mu(|F| = n) = 0$ for every fixed n, while $\mu(\exists n(|F| = n)) = 1$.

2) These are *labelled* probabilities, as the probability space consists of labelled frames, i.e. two isomorphic frames with different labelling of the domains are regarded different. Alternatively, one can consider probability spaces consisting of *unlabelled* frames, i.e. isomorphism types of frames. Though more natural from mathematical viewpoint, this choice makes the computation of μ_n much more difficult. It turns out, however, that *the labelled and unlabelled asymptotic probabilities coincide*. The reason for this is that the property of a frame to be *rigid*, i.e. not to have non-trivial automorphisms, is almost surely valid (see [Gurevich 92]). Note that every rigid n-element frame has the same number, viz. $n!$, of non-isomorphic labellings, whence the equality of the asymptotic probabilities.

The construction of random frames by means of a random pairwise assignment of a binary relation with a given probability for truth of the relation p can be performed on *infinite* sets, too. The outcome of such a random construction on the set N of natural numbers will be called a **countable random frame**. An alternative, and more precise definition will be given below, from which it will also become clear that such a countable random frame is unique up to isomorphism.

Using combinatorial-probabilistic argument, it can be proved (see [Fagin 76] or [Gurevich 92]) that every countable random frame with probability 1 satisfies an infinite set of **extension axioms**. These are all instances of the scheme **(EXT)**:

$$\forall \overline{x} \exists y \left(\bigwedge_{i \neq j} x_i \neq x_j \to \left(\bigwedge_{i \in [n]} x_i \neq y \wedge T(y, y) \wedge \bigwedge_{i \in I} Rx_i y \; \wedge \bigwedge_{i \in [n]-I} \neg Rx_i y \; \wedge \right.\right.$$

$$\left.\left. \bigwedge_{i \in J} Ryx_j \wedge \bigwedge_{j \in [n]-J} \neg Ryx_j \right)\right),$$

where $\overline{x} = x_1 \ldots x_n$, $T(y, y)$ is either Ryy or $\neg Ryy$, $[n] = \{1, 2, \ldots, n\}$ and $I, J \subseteq [n]$.

By θ_k we denote the conjunction of all (finitely many) extension axioms using at most k variables.

As a particular case of Gaifman's results (see [Gaifman 64]) the following holds.

Theorem 1. *The theory EXT is consistent and ω- categorical, hence complete.*

The unique countable model F^r of EXT will be called **the countable random frame**.

Theorem 2. ([Fagin 76]) *For any first-order sentence ψ of the first-order language for frames:*

1. *If $F^r \models \psi$ then $\mu(\psi) = 1$.*
2. *If $F^r \not\models \psi$ then $\mu(\psi) = 0$.*

This theorem immediately implies the *0-1 law for first-order logic: every first-order sentence for frames is either almost surely valid or almost surely invalid* in the finite. Moreover, it implies, by compactness, that *every almost surely valid first-order sentence follows from finitely many extension axioms*, hence from some θ_k.

The extension axioms can be similarly defined in any relational language and Gaifman's result applies and determines the *countable random structure* of that language. Furthermore, Fagin's theorem holds for the first-order logic of any relational language.

2.3. *On the almost sure model validity of modal formulas*

One can consider probability spaces of *random Kripke models*, i.e. models based on random frames and with random valuation of the propositional variables.

Following [Halpern and Kapron 94] we denote the corresponding probabilities by $\nu_{n,p}(\phi)$ and $\nu_p(\phi)$ and will call validity in models *structure validity*.

Since the evaluated propositional variables can be regarded as unary predicates, Kripke models are first-order relational structures and the 0-1 law holds for the almost sure structure validity. The remarks on probabilities on frames as first-order structures apply here, too, but there is an added subtlety: the probability space of random models depends on the set $\mathcal{P}$ of propositional variables, too, and therefore one can expect that the probability of a formula can depend on the cardinality of $\mathcal{P}$. As shown in [Halpern and Kapron 94], however, this is not the case, so without loss of generality, when computing asymptotic probabilities one needs to consider only languages with finite sets $\mathcal{P}$.

The following results about almost sure structure validity were obtained in [Halpern and Kapron 94].

Proposition 2. *ϕ is a consistent propositional formula iff $\Diamond\phi$ is valid in almost all models.*

It turns out that all formulas of the type $\Diamond\phi$ where ϕ is a consistent propositional formula axiomatize the modal logic of almost sure model validity. It turns out to be the so called *Carnap's logic*. Note that it is not a normal modal logic, since it is not closed under uniform substitutions.

Furthermore, [Halpern and Kapron 94] provides a translation from an arbitrary modal formula ϕ to a propositional formula ϕ^r, with the property that $\phi \Leftrightarrow \phi^r$ is almost-surely valid. That translation is defined by structural induction on the formulas:

- $p^r = p$ for a primitive proposition p
- $(\phi \wedge \psi)^r = \phi^r \wedge \psi^r$
- $(\neg\phi)^r = \neg\phi^r$
- $(\Box\phi)^r = \begin{cases} \top \text{ if } \phi^r \text{ is valid} \\ \bot \text{ otherwise.} \end{cases}$

Proposition 3. *The formula $\phi \Leftrightarrow \phi^r$ is valid in almost all structures.*

Theorem 3. *For every modal formula ϕ, $\nu(\phi) = 1$ iff the propositional formula ϕ^r is valid; otherwise $\nu(\phi) = 0$.*

2.4. *Almost sure frame validity*

In this paper we study almost sure validity of modal formulas in *Kripke frames*. Recall that a modal formula ϕ is **valid in a Kripke frame** F if it is valid in every model over F, i.e. for every valuation in F of the propositional variables occurring in ϕ. Thus, frame validity is a *second-order* property. More specifically, it is a monadic Π^1_1-property, via the *standard translation* of modal formulas (see [van Benthem 85]).

Again following [Halpern and Kapron 94] we denote the respective probabilities by $\mu_{n,p}(\phi)$ and $\mu_p(\phi)$. (This should not cause confusion with first-order validity in frames, as it will be clear from the context whether the formula in question in modal or first-order.)

Hereafter we will refer to (almost sure) frame validity as just (almost sure) validity. Formally:

Definition 3. *A modal formula ϕ is said to be* **almost surely valid (in the finite)** *if $\mu(\phi) = 1$, and respectively, almost surely invalid if $\mu(\phi) = 0$.*

The 0-1 law for first-order logic does not apply to almost sure frame validity. Yet:

Theorem 4. ([Halpern and Kapron 94]) *Zero-one law for frame validity in modal logic) For every modal formula ϕ, $\mu(\phi) = 1$ or $\mu(\phi) = 0$.*[2]

3. The modal logic $\mathbf{ML}^r$ of the countable random frame and the modal logic $\mathbf{ML}^{as}$ of the almost sure frame validity in the finite

3.1. Some properties of the countable random frame and of the almost sure frame validity in the finite

We begin with some basic facts about F^r, which easily follow from the extension axioms:

- It has a diameter 2, i.e. every point can be reached from any point (including itself) in 2 R-steps. Indeed, by θ_3: $F^r \models \forall x \forall y \exists z (Rxz \wedge Rzy)$.
- Every point has infinitely many predecessors and infinitely many successors.
- Every finite frame is embeddable as a subframe in F^r.

We denote the **universal modality**, interpreted by the Cartesian square of the domain, by **A** and its dual, the existential modality, by **E**. (For reference on modal logic with universal modality see [Goranko and Passy 92].) These modalities turn out to be definable *almost everywhere*:

Proposition 4. *In every frame with diameter 2:*

$$\mathbf{E}p \equiv \Diamond\Diamond p, \quad \textit{respectively} \quad \mathbf{A}p \equiv \Box\Box p,$$

Therefore this equivalence holds in almost every finite frame and in F^r.

Proposition 5. ([Kolaitis and Vardi 90]) *A Π^1_1-sentence ϕ is true in F^r iff it follows from some extension axiom, and hence every such sentence is almost surely valid (in the finite). Respectively, every not almost surely false Σ^1_1-sentence is true in F^r.*

[2] See the footnote in the introduction.

Proof. One direction is trivial, since every extension axiom is true in F^r. For the other, let $\phi = \forall\overline{\mathbf{R}}\psi$, where ψ is a first-order sentence. Suppose ϕ does not follow from any extension axiom. Then every finite set of extension axioms is satisfiable together with $\neg\psi$, hence by compactness, (EXT) is satisfiable together with $\neg\psi$, hence $\neg\psi$ is true in some $\overline{\mathbf{R}}$-expansion of F^r – a contradiction. □

Proposition 6. *Every almost surely valid modal formula A that defines a universal frame-condition is valid.*

Proof. Suppose $\neg A$ is satisfiable. Then it is satisfiable in a model on a finite frame F. The satisfiability of $\neg A$ is an existential condition, hence preserved in extensions. By the extension axioms, almost every finite frame contains a copy of F, and hence satisfies $\neg A$. Thus, A is almost surely non-valid. □

Thus, reflexivity, symmetry, transitivity etc. are not almost surely valid, hence they are almost surely invalid.

3.2. The modal logics $\mathbf{ML}^r$ *and* $\mathbf{ML}^{as}$

Definition 4.

$\mathbf{ML}^r$ *is the modal logic of the formulas valid in* F^r.

$\mathbf{ML}^{as}$ *is the modal logic of the formulas which are almost surely valid in the finite.*

Proposition 7.

1. $\mathbf{ML}^r$ *and* $\mathbf{ML}^{as}$ *are normal modal logics.*
2. $\mathbf{ML}^r \subseteq \mathbf{ML}^{as}$.
3. *A formula is in* $\mathbf{ML}^r$ *iff it follows logically from an extension axiom (meaning that it is valid in every frame which, as a first-order structure, is a model of that extension axiom).*

Proof. (1) A routine verification shows that both logics contain all tautologies, the axiom K, and are closed under substitution, MP and necessitation. (2) and (3) follow from proposition 5. □

3.3. Some almost surely valid modal principles

Here are some validities in ML^r that follow from the extension axioms.

Proposition 8. *The following modal formulas follow from (EXT) and hence are valid in* F^r.

- $\Box p \to \Diamond p$, *and hence* $\Diamond\top$ *(seriality);*
- $\Diamond\Box p \to \Diamond p$, $\Box p \to \Box\Diamond p$;
- $p \to \Diamond\Diamond p$, $\Box\Box p \to p$ *(diameter 2);*
- $\Diamond p \to \Diamond\Diamond p$, $\Box\Box p \to \Box p$ *(density);*
- $\Diamond\Box p \to \Box\Box\Diamond p$, $\Diamond\Diamond\Box p \to \Box\Diamond p$.

Note that these imply *Church-Rosser property:* $\Diamond\Box p \to \Box\Diamond p$.

- $\mathbf{M}\Diamond\Diamond p \equiv \Diamond\Diamond p$, $\mathbf{M}\Box\Box p \equiv \Box\Box p$ *for any string* $\mathbf{M}$ *of modalities.*

Two useful derivable rules of inference in ML^r:

- If $\vdash \Box\phi$ then $\vdash \phi$.
- If $\vdash \phi$ then $\vdash \Diamond\phi$.

Proposition 9. *For any strings of modalities* $\mathbf{M_1}$ *and* $\mathbf{M_2}$*, the modal reduction principle*

$\mathbf{M_1}\Diamond p \to \mathbf{M_2}\Box p$ *is almost surely structure-invalid and hence almost surely frame-invalid.*

Proof. $(\mathbf{M_1}\Diamond p \to \mathbf{M_2}\Box p)^r = \bot$. □

For instance, McKinsey's axiom $\Box\Diamond p \to \Diamond\Box p$ is almost surely frame-invalid.

3.4. Depth 3 equivalence and almost sure isomorphism

Lemma 1. *Let* $F = \langle W, R\rangle$ *be a finite frame,* $W = \{w_1, \ldots, w_n\}$ *and let* χ_F *be the formula*

$$\neg\mathbf{A}\Big(\bigvee_{i=1}^{n}\mathbf{E}p_i \wedge \bigwedge_{1\le i\ne j\le n}(p_i \to \neg p_j) \wedge \bigwedge_{1\le i,j\le n}\{p_i \to \Diamond p_j | w_i R w_j\}\wedge$$

$$\bigwedge_{1\le i,j\le n}\{p_i \to \neg\Diamond p_j | \neg w_i R w_j\}\Big),$$

for different propositional variables $\{p_1, \ldots, p_n\}$.

Then for every frame G *with diameter 2,* F *is a p-morphic image of* G *iff* $G \not\models \chi_F$.

Proof. Suppose $G, V \not\models \chi_F$ for some valuation V. Then every point $y \in G$ satisfies exactly one variable $p_{i(y)}$ from $\{p_1, \ldots, p_n\}$. Furthermore, the mapping $f : G \longrightarrow F$ defined by $f(y) = w_{i(y)}$ is a surjective p- morphism.

Vice versa, if $f : G \longrightarrow F$ is a surjective p-morphism, then the valuation V on G defined by $V(p_i) = f^{-1}(w_i)$ satisfies $\neg\chi_F$. □

Proposition 10. *Let for* $n \in N$, $P_n = Pr(F \cong G | F \equiv_{3,n} G)$ *be the probability that two randomly chosen frames* F, G *of size at most n are isomorphic given that they validate the same modal formulas of at most n propositional variables and modal depth 3.*

Then $\lim_{n\to\infty} P_n = 1$.

Proof. Suppose that $F = \langle W, R\rangle$ and $G = \langle U, S\rangle$ validate exactly the same depth 3 formulas. Then, in particular $G \not\models \chi_F$ and $F \not\models \chi_G$. With asymptotic probability 1 both frames have diameter 2. Then, by lemma 1 each of them is a p-morphic image of the other, so they are isomorphic. □

4. Complete axiomatization of $\mathbf{ML}^r$

4.1. Axioms for $\mathbf{ML}^r$

What properties of F^r can be defined modally? Certainly those must be preserved by the basic constructions preserving modal validity. Since generated subframes and disjoint unions are trivialized in a language with universal modality, only p-morphisms can be meaningfully used, and the following lemma gives a clue.

Lemma 2. *(See [van Benthem 85], Th 15.11) A first-order sentence in the language with* $=$ *and a binary relation* R *is preserved by p-morphisms iff it is equivalent to one constructed from atomic formulas,* $\top$*, and* $\bot$ *using* $\wedge, \vee, \exists, \forall$*, and restricted universal quantification* $\forall z(Ryz \to \ldots)$ *for* $z \neq y$.

Therefore, the instances of the following schema are preserved by p-morphisms:

$$\forall x_1 \ldots \forall x_n \exists y (\bigwedge_{i \in I} Rx_i y \wedge \bigwedge_{j \in J} Ryx_j),$$

where $I, J \subseteq [n] = \{1, \ldots, n\}$.

This is an approximation of the schema EXT. Its strongest form is EXT_p:

$$\forall x_1 \ldots \forall x_n \exists y \bigwedge_{i \in [n]} (Rx_i y \wedge Ryx_j).$$

As we will see, this is all that modal logic can express from the extension axioms.

We now introduce the following *system of axioms for* $\mathbf{ML}^r$ (recall that $\mathbf{A}$ is defined as $\Box\Box$).

(ML^r_1) K: $\Box(p \to q) \to (\Box p \to \Box q)$.
(ML^r_2) $\mathbf{A}p \to p$.
(ML^r_3) $\mathbf{A}p \to \mathbf{A}\Box p$.
(ML^r_4) $p \to \mathbf{AE}p$.
(ML^r_5) The scheme MODEXT consisting of the sequence of axioms MODEXT_n for $n \in N$:

$$\bigwedge_{k=1}^{n} \mathbf{E}(p_k \wedge \Box q_k) \to \mathbf{E} \bigwedge_{k=1}^{n} (\Diamond p_k \wedge q_k).$$

Theorem 5. *The axiomatization of* $\mathbf{ML}^r$ *is sound.*

Proof. Straightforward. Note that MODEXT is a modal translation of the scheme EXT_p. □

Let us denote the canonical model for $\mathbf{ML}^r$ by M^c and its underlying frame by $F^c = \langle W^c, R^c \rangle$.

For the completeness proof we will need the following technical result.

Lemma 3. *Let A, B be finite subsets of F^c. Then there is a reflexive point $y \in F^c$ such that $y \notin A$ and for every $x \in B$, both $R^c xy$ and $R^c yx$ hold.*

Proof. First, we show that the rôle of A is inessential. Indeed, for every $a \in A$ there is $y_a \in M^c$ such that $\neg R y_a a$: for any propositional variable p, either $p \notin a$ or $\neg p \notin a$, while both $\Box\neg p$ and $\Box p$ are consistent. If, for instance $p \notin a$ then we choose y_a to contain $\Box\neg p$, otherwise we choose it to contain $\Box p$. So, we can add to B such y_a for every $a \in A$. Then any y satisfying the requirement of the lemma will be different from any $a \in A$. Hence, it suffices to prove the claim for $A = \emptyset$.

Let $B = \{b_1, \dots b_n\}$. Some notation: for any $x \in M^c$, $\Box x = \{\alpha | \Box\alpha \in x\}$, and $\Diamond(x) = \{\Diamond\alpha | \alpha \in x\}$.

The key fact to prove is that the set

$$Y = \Box b_1 \cup \dots \cup \Box b_n \cup \Diamond(b_1) \cup \dots \cup \Diamond(b_n) \cup \{\Box\gamma \to \gamma | \text{ for all formulas } \gamma\}$$

is $\mathbf{ML}^r$-consistent.

Suppose otherwise. Then some finite subset Y_0 of Y is inconsistent. Since all sets $\Box x$ are closed under conjunctions, we can assume that Y_0 consists of:

$\alpha_1 \in \Box b_1, \dots, \alpha_n \in \Box b_n, \Diamond\beta_{11}, \dots, \Diamond\beta_{1k_1} \in \Diamond(b_1), \dots, \Diamond\beta_{n1}, \dots, \Diamond\beta_{nk_n} \in \Diamond(b_n)$,

and $\Box\gamma_1 \to \gamma_1, \dots, \Box\gamma_m \to \gamma_m$.

Let $\beta_i = \beta_{i1} \wedge \dots \wedge \beta_{ik_i}$. Then $\beta_i \in b_i$, for $i = 1, ..., n$. Let us denote derivability in $\mathbf{ML}^r$ by $\vdash$.

Since Y_0 is inconsistent,

$$\vdash \bigwedge_{i\in[n]} \alpha_i \wedge \bigwedge_{i\in[n], j\in[k_i]} \Diamond\beta_{ij} \to \neg((\Box\gamma_1 \to \gamma_1) \wedge \dots \wedge (\Box\gamma_m \to \gamma_m)),$$

i.e.

$$\bigwedge_{i\in[n]} (\alpha_i \wedge \bigwedge_{j\in[k_i]} \Diamond\beta_{ij}) \vdash (\Box\gamma_1 \wedge \neg\gamma_1) \vee \dots \vee (\Box\gamma_m \wedge \neg\gamma_m),$$

hence

$$\bigwedge_{i\in[n]} (\alpha_i \wedge \Diamond\beta_i) \vdash (\Box\gamma_1 \wedge \neg\gamma_1) \vee \dots \vee (\Box\gamma_m \wedge \neg\gamma_m).$$

Therefore

$$\mathbf{E}(\bigwedge_{i\in[n]} (\alpha_i \wedge \Diamond\beta_i)) \vdash \mathbf{E}((\Box\gamma_1 \wedge \neg\gamma_1) \vee \dots \vee (\Box\gamma_m \wedge \neg\gamma_m)),$$

hence by MODEXT:

$$\bigwedge_{i\in[n]} \mathbf{E}(\Box\alpha_i \wedge \beta_i) \vdash \mathbf{E}((\Box\gamma_1 \wedge \neg\gamma_1) \vee \dots \vee (\Box\gamma_m \wedge \neg\gamma_m)).$$

But, for every i, $(\Box\alpha_i \wedge \beta_i) \in b_i$, hence $\mathbf{E}(\Box\alpha_i \wedge \beta_i)$ is valid in M^c, and therefore derivable in $\mathbf{ML}^r$. Then $\mathbf{E}((\Box\gamma_1 \wedge \neg\gamma_1) \vee \dots \vee (\Box\gamma_m \wedge \neg\gamma_m))$ is derivable in $\mathbf{ML}^r$ which contradicts the soundness of $\mathbf{ML}^r$ for F^r because this formula is almost surely structure-invalid, hence almost surely frame-invalid, so not valid in F^r.

Thus, Y must be consistent. Let y be any maximal theory in $\mathbf{ML}^r$ which extends Y. Then $\Box b_i \subseteq y$, so $R^c b_i y$, and $\Diamond(b_i) \subseteq y$, so $R^c y b_i$, for each $i = 1, ..., n$. Finally, $\Box y \subseteq y$, i.e. $R^c yy$. □

Theorem 6. *The axiomatization of* $\mathbf{ML}^r$ *is complete.*

Proof. First, note that the axioms derive $\mathbf{A}p \to \Box p$ and $\mathbf{A}p \to \mathbf{AA}p$, so they ensure that $\mathbf{A}$ is an S5-modality implying $\Box$. In fact, this renders a complete axiomatization of the universal modality added to $\Box$, as proved in [Goranko and Passy 92].

To be precise, the relation R^u corresponding to $\mathbf{A}$ in M^c is not necessarily the universal relation, but an equivalence relation containing R^c. However, every R^u-cluster is a generated submodel of R^c and R^u is the universal relation there. For the purpose of proving completeness we only need to deal with generated submodels of M^c, so we can assume without further ado that $\mathbf{A}$ *is* the universal modality on M^c.

Furthermore, all axioms are Sahlqvist formulas, hence the logic is canonical.

We need to show that every modal formula ϕ satisfiable in M^c is satisfiable in some model over F^r.

Hereafter we restrict the language to those propositional variables that occur in ϕ.

Let k be the modal depth of ϕ and suppose $M^c, x_0 \models \phi$.

The idea in a nutshell: We are going to build up step by step a model M^r over F^r, a submodel M_ω of M^c, and a p- morphism from M^r onto M_ω. Besides, (M_ω, x_0) will be k-bisimilar to (M^c, x_0), hence $M_\omega, x_0 \models \phi$, so ϕ will be satisfied in M^r.

Now, the technicalities.

First, enumerate $F^r : s_0, s_1, ...$ choosing s_0 to be reflexive iff x_0 is reflexive.

We shall define inductively on n two chains of finite models:
$M_0^c \subset M_1^c \subset \ldots \subset M^c$, based on frames $F_0^c \subset F_1^c \subset \ldots \subset F^c$, and
$M_0^r \subset M_1^r \subset \ldots$, based on frames $F_0^r \subset F_1^r \subset \ldots \subset F^r$,
and a chain of functions $f_o \subset f_1 \subset \ldots$ where $f_n : F_n^r \longrightarrow F_n^c$, such that:

(1) $x_0 \in M_0^c$ and $s_0 \in M_0^r$;
(2) f_n is a p-morphism of M_n^r onto M_n^c;
(3) $M^r = \bigcup_{n\in\omega} M_n^r$ covers F^r;
(4) $M_\omega = \bigcup_{n\in\omega} M_n^c$ is a submodel of M^c such that $(M_\omega, x_0) \sim_k (M^c, x_0)$.

Before we proceed with the construction, let us note that (2) implies that $f = \bigcup_{n\in\omega} f_n$ is a p-morphism from M^r onto M_ω, so the claim will follow.

Now, the inductive construction.

For $n = 0$ put $F_0^c = \{x_0\}$, $F_0^r = \{s_0\}$, $f_0(s_0) = x_0$, and the valuation at s_0 to match the valuation at x_0 in M^c.

For the inductive step, suppose M_n^c, M_n^r, and f_n are defined as required. We will break the construction of M_{n+1}^c, M_{n+1}^r, and f_{n+1} down to several steps.

i) This step will ensure that eventually M^r covers the whole F^r.

Let t_n be the first point in $s_0, s_1, \ldots$ which is not in F_n^r. Let $P(t_n)$ and $S(t_n)$ be respectively the set of predecessors and the set of successors of t_n in F_n^r, and let $P' = f_n[P(t_n)]$, $S' = f_n[S(t_n)]$.

ii) Select a point y_n from $M^c - M_n^c$ to "simulate" the position of t_n in M_n^r, i.e. such that all points in P' are predecessors of y_n and all points in S' are successors of y_n. The axiom scheme MODEXT, strengthened by lemma 3, ensures that such y_n exists. Furthermore, if t_n is reflexive we require that y_n is reflexive, too.

iii) Now, there are 4 possible types of defects to fix:

(a) Unforeseen R^c-arrows coming to y_n from outside P'. This also covers cases where $f_n(s') = f_n(s'')$ and $R^r s' t_n$ but not $R^r s'' t_n$.
(b) In particular, y_n can be reflexive, while t_n is not.
(c) Unforeseen R^c-arrows going from y_n to a set of successors $S_0(y_n)$ outside S'.
(d) The points in $F_n^c - F_{n-1}^c$, i.e. those added at the previous inductive step of the construction, and y_n itself, do not have all necessary successors to extend the k-bisimulation with M^c.

To fix all these defects we will extend M_n^c and M_n^r with more points as follows.

To fix (a) and (b): Select a new point $u_n \in F^r$ to match the in- and out- R^c-arrows of y_n in F_n^c. Furthermore, if (b) is the case, require that each of t_n and u_n is a successor of the other.

Now, extend f_n with $f_n(t_n) = f_n(u_n) = y_n$.

To fix (c) and (d): the set of 'unsaturated' points at the moment is $X_n = \{y_n\} \cup (F_n^c - F_{n-1}^c)$ (where $F_{-1}^c = \emptyset$).

For each $x \in X_n$ let $S^k(x)$ be a set of successors of x in $M^c - M_n^c$ containing one representative of each modal $(k-1)$-type realized by a successor of x in M^c and still missing in M_n^c.

Let $U_n^c = S_0(y_n) \cup \bigcup_{x \in X_n} S^k(x)$. Note that U_n^c is finite.

Now, we define

$$M_{n+1}^c = M_n^c \cup U_n^c \cup \{y_n\}.$$

iv) We need to add p-morphic pre-images for U_n^c to M_n^r in order to extend the p-morphism f_n over M_{n+1}^c.

Let U_n^r be a subset of $F^r - F_n^r$ such that there is a frame isomorphism $h : U_n^c \longrightarrow U_n^r$ and moreover U_n^r matches the configuration of U_n^c in M_{n+1}^c, i.e:

• For every $y \in U_n^c$ and $t \in F_n^r \cup \{t_n, u_n\}$, $h(y)$ is a predecessor (successor) of t iff y is a predecessor (successor) of $f_n(t)$. In particular, every point in $h(S_0(y_n))$ is a successor of both t_n and u_n.

That will fix the defects (c) and (d).

Note that in order to select the set U_n^r we make use of an extension axiom for F^r with unbounded number of variables.

v) Finally, we define:

$$F^r_{n+1} = F^r_n \cup U^r_n \cup \{t_n, u_n\},$$
$$f_{n+1} = f_n \cup h^{-1} \cup \{(t_n, y_n), (u_n, y_n)\},$$

and extend M^r_n to M^r_{n+1} accordingly to extend the p- morphism from f_n to f_{n+1}.

This completes the inductive construction.

Note that the condition (2) for f_{n+1} is guaranteed by the inductive hypothesis and the construction.

Now, we take the unions:

$$M^r = \bigcup_{n\in\omega} M^r_n,$$

$$M_\omega = \bigcup_{n\in\omega} M^c_n.$$

Condition (3) is immediate from the construction.

It remains to prove (4). We shall apply prop. 1. The key observation is that every $x \in M_\omega$ has the same k- type in M_ω as in M^c. In order to prove this, we show by an easy induction on $m \leq k$ that if $x \in M_n$ then x has the same l-type in M_{n+m} as in M^c, using again prop. 1. □

Proposition 11. $\mathbf{ML}^r$ *is not finitely axiomatizable.*

Proof. According to a well-known criterion for non-finite axiomatizability due to Tarski (see e.g. [Chagrov and Zakharyaschev 97]) it is sufficient to construct a strictly increasing sequence of modal logics $\mathbf{L}_1 \subset \mathbf{L}_2 \subset ...$ such that $\bigcup_{k=1}^{\infty} \mathbf{L}_k = \mathbf{ML}^r$.

Let for any $k \in N$, $\phi_k = \mathrm{MODEXT}_k$.

Further, let $\mathbf{ML}^r_k$ be the sublogic of $\mathbf{ML}^r$ axiomatized by the modal formulas which follow from θ_k. Then we define the sequence $\{\mathbf{L}_k\}_{k\in N}$ as a subsequence of $\{\mathbf{ML}^r_k\}_{k\in N}$ inductively as follows:

- $\mathbf{L}_1 = \mathbf{ML}^r_1$.
- Suppose $\mathbf{L}_k = \mathbf{ML}^r_{n_k}$ for some $n_k \in N$. Then we define $\mathbf{L}_{k+1} = \mathbf{ML}^r_{n_{k+1}}$ where n_{k+1} is the least positive integer such that $\phi_{n_{k+1}-1}$ is not valid in some model of θ_{n_k}. To see that such an index exists, take any model of θ_{n_k}, say of size m. Then no point in that model is related to every point in the model, so ϕ_m fails there for $p_1, \ldots, p_m$ evaluated in pairwise different singleton sets.

Thus, $\mathbf{L}_k \not\vdash \phi_{n_{k+1}-1}$, while $\theta_{n_{k+1}} \models \phi_{n_{k+1}-1}$, so $\mathbf{L}_{k+1} \vdash \phi_{n_{k+1}-1}$, hence $\mathbf{L}_{k+1}$ is a proper extension of $\mathbf{L}_k$.

Finally, $\bigcup_{k=1}^{\infty} \mathbf{L}_k = \bigcup_{k=1}^{\infty} \mathbf{ML}^r_k = \mathbf{ML}^r$. □

This proof can be simplified when we describe the finite frames of ML^r below.

4.2. The finite frames and the finite frames property of ML^r

Definition 5. *Let $F = \langle W, R\rangle$ and $x \in W$. x is a* **central point** *in F if Rxy and Ryx hold for every $y \in W$.*

Proposition 12. *For every finite frame F the following are equivalent.*

1. $F \models ML^r$.
2. F has a central point.
3. F is a p-morphic image of F^r.

Proof. (1) $\Leftrightarrow$ (2): Every finite frame of size n which validates MODEXT$_n$ has a central point and every finite frame with a central point validates MODEXT.

(3) $\Rightarrow$ (1): p-morphisms preserve modal validity.

(1) $\Rightarrow$ (3): Suppose otherwise. Then $F^r \models \chi_F$, hence $\chi_F \in ML^r$, so $F \models \chi_F$ - a contradiction. □

Proposition 13. *Every finite frame F obtained from F^r by filtration has a central point, and hence validates MODEXT.*

Proof. It is sufficient to prove the claim for the case when $F = \langle W, R\rangle$ is obtained from F^r by a minimal filtration f. Let $W = \{w_1, \ldots, w_n\}$ and $\{x_1, \ldots, x_n\}$ be points in F^r such that $f(x_i) = w_i$, $i = 1, \ldots, n$. Then there is $y \in F^r$ such that $R^r yx_i$ and $R^r x_i y$ for each $i = 1, \ldots, n$. Therefore, $f(y)$ is a central point in F. □

Corollary 1. *The logic ML^r has the finite frame property and the complexity of its satisfaction problem is in EXPTIME.*

Proof. Every ML^r-consistent formula ϕ is satisfiable in F^r, hence it is satisfiable in a finite frame of size at most $2^{|\phi|}$ obtained from F^r by filtration. By proposition 13 every such frame validates ML^r. Furthermore, these frames are recognizable directly, without checking validity of the axioms in them, in polynomial time. Now, the proof follows the one for K plus universal modality (see [Spaan 93]). □

The proofs of all results in this section easily extend to temporal logic with minor modifications required by the temporal versions of the notions of p-morphism, bisimulation and filtration. Thus, we obtain:

Proposition 14. *1. The axioms of ML^r added to the minimal temporal logic T provide a sound and complete axiomatization of the temporal logic TL^r of F^r.*
2. The finite frames of TL^r are precisely the frames with central point.
3. TL^r has the finite frame property and is decidable in EXPTIME.

4.3. *Kernels in random finite frames and in F^r*

Every p-morphic image G of a frame F determines a kernel partition in F. As we have shown in lemma 1, the existence of the kernel partition determined by G in a random frame is characterized by the non-validity of the respective χ_G in that frame. On the other hand, the asymptotic probability of such a partition can be estimated using well-developed combinatorial-probabilistic methods. These relationships can be used to obtain results on almost sure validity and invalidity of various modal formulas. Here we shall consider two very simple examples, and one of them will turn out to distinguish ML^r from ML^{as}.

Definition 6. *Let $F = \langle W, R \rangle$ be a frame, and A be a non-empty subset of W. A is called:*

- **independent** *if $F \models \forall x, y \in A(\neg Rxy)$.*
- **dominating over a subset** *B of F if $F \models \forall x(x \in B \to \exists y \in A(Rxy))$.*
- **dominating set** *if it is dominating over $W - A$.*
- **kernel** *if it is both independent and dominating.*
- **double kernel** *if it is a kernel and can be split in two non-empty 'sub-kernels' A_1 and A_2 such that $F \models \forall x(x \notin A \to (\exists y \in A_1(Rxy) \wedge \exists y \in A_2(Rxy))$.*

The notion of a kernel has been independently studied in the theory of digraphs, i.e. loopless frames. It was proved in [Fernandez de la Vega 90] that *almost every finite digraph has a kernel*. Since existence of a kernel is a Σ^1_1-property, it follows that the countable random digraph has a kernel, too. These results, however, need not hold for frames because *almost every frame is not a digraph*, i.e. the asymptotic probability of a frame to be a digraph is 0, and therefore no results about asymptotic probabilities on digraphs transfer directly to frames.

Consider the frames
$\mathbf{K}_2 = \langle \{x, y\}, \{(x, x), (x, y), (y, x)\} \rangle$ and
$\mathbf{K}_3 = \langle \{x, y_1, y_2\}, \{(x, x), (x, y_1), (x, y_2), (y_1, x)(y_2, x)\} \rangle$.
Note that K_2 is a p-morphic image of K_3.

Proposition 15. *Let F be a frame such that $F \models \theta_3$. Then:*

1. *F has a kernel iff $\mathbf{K}_2$ is a p-morphic image of F.*
2. *F has a double kernel iff $\mathbf{K}_3$ is a p-morphic image of F.*
3. *In particular, F^r has a double kernel.*

Proof. We will prove (2); (1) is similar, but simpler. If F has a double kernel, then the function $f : F \longrightarrow \mathbf{K}_3$ which sends the points from the two sub-kernels of the double kernel resp. to y_1 and y_2 and the rest to x is a p-morphism. To check this it is sufficient to use the extension axioms in θ_3. (In particular, θ_3 ensures that every point in the kernel sends an arrow to the complement of the kernel.) Conversely, if f is a p-morphism of F onto $\mathbf{K}_3$, then $f^{-1}(y_1)$ and $f^{-1}(y_2)$ are the sub-kernels of a double kernel in F.

Finally, since $\mathbf{K}_3$ has a central point x, it is a p-morphic image of F^r, by lemma 12. □

Thus, the formula χ_{K_2} (resp. χ_{K_3}) is valid precisely in those frames satisfying θ_3 which do not have kernels (resp. double kernels). Here are slightly simpler formulas with the same properties:

NO-KER $= \mathbf{E}(p \leftrightarrow \Diamond p)$, and

NO-DKER $= \mathbf{E}((p \vee q) \wedge \Diamond(p \vee q)) \vee \mathbf{E}(\neg(p \vee q) \wedge (\Box\neg p \vee \Box\neg q))$.

Proposition 16. *For every frame F satisfying θ_3:*

1. *F has a kernel iff $F \not\models$ NO-KER.*
2. *F has a double kernel iff $F \not\models$ NO-DKER.*

Proof. NO-KER is falsified by a model on F iff the valuation of p is a kernel in F. Likewise, NO-DKER is falsified by a model on F iff the valuations of p and q are the sub-kernels of a double kernel in F. □

Using kernels one can show the invalidity in F^r of various formulas which are not first-order definable. Here is an example.

Proposition 17. *Each of the following modal reduction principles, with alternating modalities in the antecedent and consequent, fails in F^r:*

- $\Diamond\Box \ldots p \to \Box\Diamond \ldots \Box p$.
- $\Diamond\Box \ldots \Diamond p \to \Box\Diamond \ldots p$.
- $\Box\Diamond \ldots \Box\Diamond p \to \Diamond\Box \ldots p$.
- $\Diamond\Box \ldots p \to \Diamond\Box \ldots \Diamond\Box p$.

Proof. Each of these is falsified when p is evaluated either in the kernel or in its complement.[3] □

Theorem 7. *The existence of a double kernel is almost surely false in the finite.*

Proof. We begin with combinatorial-probabilistic calculation of the expectation of the number DK^n of double kernels in a random frame from $\mathcal{S}(n, p)$. Put $q = 1 - p$.

Let $G \in \mathcal{S}(n, p)$, $G = \langle W, R\rangle$, and Y be an m-element set in G. The probability for Y to be independent is q^{m^2}. Furthermore, for any k-element subset Z of Y, the probabilities of Z and $Y - Z$ to be dominating over $W - Y$ are respectively $(1 - q^k)^{n-m}$ and $(1 - q^{m-k})^{n-m}$. Since these are independent events, the probability of Y to be a double kernel with sub-kernels Z and $Y - Z$ is $q^{m^2}(1 - q^k)^{n-m}$ $(1 - q^{m-k})^{n-m}$. Then, the probability of Y to be a double kernel is not greater than

$$q^{m^2} \sum_{k=0}^{m} \binom{m}{k} (1 - q^k)^{n-m} (1 - q^{m-k})^{n-m}.$$

Therefore the expectation of DK^n satisfies

$$E(DK^n) \leq \sum_{m=2}^{n} \binom{n}{m} q^{m^2} \sum_{k=0}^{m} \binom{m}{k} (1 - q^k)^{n-m} (1 - q^{m-k})^{n-m}.$$

[3] We thank Jean-Marie le Bars who first applied this argument to a particular case.

Since $q^k + q^{m-k} \geq 2q^{\frac{m}{2}}$, we get $(1-q^k)(1-q^{m-k}) \leq (1-q^{\frac{m}{2}})^2$, hence

$$\sum_{k=0}^{m} \binom{m}{k} (1-q^k)^{n-m} (1-q^{m-k})^{n-m} \leq (1-q^{\frac{m}{2}})^{2(n-m)} \sum_{k=0}^{m} \binom{m}{k} 2^m (1-q^{\frac{m}{2}})^{2(n-m)}.$$

Therefore

$$E(DK^n) \leq \sum_{m=2}^{n} \binom{n}{m} q^{m^2} 2^m (1-q^{\frac{m}{2}})^{2(n-m)}.$$

We are going to show that the right hand side tends to 0 as n tends to infinity, and hence $\lim_{n\to\infty} E(DK^n) = 0$.

First, we use a well-known asymptotic estimation of the binomial coefficients (see e.g. [Palmer 85]).

$$\binom{n}{m} = \frac{O(1)}{\sqrt{2\pi m}} \left(\frac{ne}{m}\right)^m .$$

Using $(1-q^{\frac{m}{2}})^{2(n-m)} \sim \exp(-2q^{\frac{m}{2}}(n-m))$ and putting $c = q^{\frac{1}{2}}$ we obtain:

$$\begin{aligned} E(DK^n) &\leq \sum_{m=2}^{n} \frac{O(1)}{\sqrt{2\pi m}} \left(\frac{ne}{m}\right)^m c^{2m^2} 2^m \exp(-2c^m(n-m)) \\ &< O(1) \sum_{m=2}^{n} (2c^{\frac{m}{2}})^m \left(\frac{ne}{m}\right)^m c^{m^2} c^{\frac{m^2}{2}} \exp(-nc^m) \exp(-nc^m) \exp(2mc^m) \end{aligned}$$

(Here we used $c^{2m^2} = c^{\frac{m^2}{2}} c^{\frac{m^2}{2}} c^{m^2}$ and $2^m c^{\frac{m^2}{2}} = (2c^{\frac{m}{2}})^m$.)

$$= O(1) \sum_{m=2}^{n} (2c^{\frac{m}{2}})^m \left(\frac{enc^m}{m} \exp\left(\frac{-nc^m}{m}\right)\right)^m \exp(2mc^m) c^{\frac{m^2}{2}} \exp(-nc^m).$$

Now: since mc^m tends to 0, $\exp(2mc^m)$ is bounded above, so it is subsumed by $O(1)$. Likewise for $(2c^{\frac{m}{2}})^m$.

To show the same for $X = \left(\frac{enc^m}{m} \exp\left(\frac{-nc^m}{m}\right)\right)^m$ we denote[4] $\alpha = \frac{nc^m}{m}$ and obtain: $X = \left(\alpha e^{1-\alpha}\right)^m$. By standard calculus, we find that $\alpha e^{1-\alpha}$ is bounded above by 1, hence $X \leq 1$. Thus,

$$E(DK^n) < O(1) \sum_{m=2}^{n} c^{m^2} e^{-nc^m} < O(1) \sum_{m=2}^{n} c^m e^{-nc^m}$$

Finally, the latter sum is majorized by $\int_1^n c^x e^{-nc^x} dx = \ln c \, \frac{e^{-cn} - e^{-nc^n}}{n}$ which clearly tends to 0 as n tends to infinity.

Therefore $\lim_{n\to\infty} E(DK^n) = 0$, and hence $\lim_{n\to\infty} Pr(DK^n \geq 1) = 0$. □

[4] This trick we borrow from [Fernandez de la Vega 90].

Remark 2. We do not have a proof that existence of a single kernel in a random frame is not almost surely valid, but we have a very strong numerical evidence for that[5]. Numerical tests using a formula estimating the expected number of kernels in a random frame from $\mathcal{S}(n, p)$, performed for n up to 10^9 indicate that for $p = 0.5$ and large enough n $E(NK^n) < 0.15$ i.e bounded way below 1. It is notable that for $n = 10^9$ and $p = 0.5$ the expected number of kernels of size m is greater than 10^{-6} only for $24 \leq m \leq 26$.

This is a demonstration of an interesting phenomenon, indicated in [Fernandez de la Vega 90], that the expected size of the kernel concentrates around $\log_{\frac{1}{q}} n - \log_{\frac{1}{q}} \log_{\frac{1}{q}} n$ and only a handful of the closest values of m to this number produce significant terms in this sum.

Corollary 2. *The formula NO-DKER is almost surely valid in the finite, hence it is in ML^{as}, but not in ML^r.*

Corollary 3. *The formula NO-DKER does not follow from the extension axioms.*

Furthermore, consider the ML^r-frames

$\mathbf{K}_{m+1} = \langle \{x, y_1, \ldots, y_m\}, \{(x, x), (x, y_1), \ldots, (x, y_m), (y_1, x), \ldots, (y_m, x)\}\rangle$, for $m > 2$.

Note that K_3 is a p-morphic image of each of these. Therefore, every χ_{K_m} is almost surely valid in the finite, while none of them is in ML^r.

5. Concluding remarks

In this paper we have obtained explicit axiomatic and model-theoretic descriptions of the logic of the countable random frame. We believe that these results can be further extended to various enriched modal formalisms such as logics with difference modality, guarded fragments, mu-calculus, and multimodal logics such as PDL.

In the light of the results from the previous section, ML^r does not axiomatize the almost sure validity in the finite, and its complete axiomatization is still an open problem. We conjecture that this axiomatization can be obtained by adding to the axioms of ML^r all formulas χ_G for the finite frames G for the logic ML^r which are almost surely not p-morphic images of random finite frames. Thus, the Π^1_1-formulae expressing non-existence of certain kernel partitions appear to play a rôle analogous to the extension axioms for first-order logic in the essentially second-order layer of the monadic Π^1_1-logic.

Finally, the study of almost sure validities and 0-1 laws can be relativized to *classes of frames* e.g. reflexive, transitive, partial orderings, etc. Some results in that direction have been obtained in [Halpern and Kapron 94]. Furthermore, it is interesting to determine for which classes of frames there is a naturally defined countable random frame, and to investigate the relationship between its modal logic and the almost sure validity in such classes.

[5] Meanwhile, Le Bars has proved in [le Bars 02] that existence of a single kernel in a random frame is almost surely false.

Acknowledgements. We thank Moshe Vardi, Jean-Marie le Bars, Lalo Fernandez de la Vega and Thierry Lacoste for stimulating and useful discussions and correspondence on 0-1 laws and asymptotic probabilities on random structures, Minko Markov for programming some numerical tests on the expected number of kernels in random frames, and Andreas Blass, Rainer Schamm, and an anonimous referee for some corrections.

The work of Valentin Goranko was done during his sabbatical leave to the University of Victoria, BC, Canada, and supported in part by NSERC.

References

[le Bars 99a] Le Bars, J.-M.: The 0-1 law fails for monadic existential second-order logic on undirected graphs, Information Processing Letters, to appear

[le Bars 99b] Le Bars, J.-M.: Counterexamples of the 0- 1 law for fragments of existential second-order logic: an overview, Bulletin of Symbolic Logic, **6**, 67–82 (2000)

[le Bars 02] Le Bars, J.-M.: The 0-1 law fails for frame satisfiability of propositional modal logic, Proc. of LICS'02, IEEE Computer Society Press, 2002

[van Benthem 85] van Benthem, J.: *Modal Logic and Classical Logic*, Bibliopolis, Napoli, 1985

[Blackburn, de Rijke, and Venema 2001] Blackburn, P., de Rijke, M., Venema, Y.: *Modal Logic*, Cambridge Univ. Press (2001)

[Blass, Gurevich and Kozen 85] Blass, A., Gurevich, Y., Kozen, D.: A zero-one law for a logic with a fixed-point operator, Information and Control, **67**, 70–90 (1985)

[Blass, Harary 79] Blass, A., Harary, F.: Properties of Almost all Graphs and Complexes, J. Graph theory, **3**, 225–240 (1979)

[Carnap 50] Carnap, R.: *Logical Foundations of Probability*, Chicago: Univ. of Chicago Press, 1950

[Chagrov and Zakharyaschev 97] Chagrov, A., Zakharyaschev, M.: *Modal Logic*, Oxford: Clarendon Press, 1997

[Compton 88] Compton, K.J.: 0-1 laws in logic and combinatorics, In: *NATO Adv. Study Inst. on Algorithms and Order*, I. Rival (ed.), Reidel, 353–383 (1988)

[Dawar and Grädel 95] Dawar, A., Grädel, E.: Generalized quantifiers and 0-1 laws, *Proc. 10th IEEE Symp. on Logic in Computer Science*, IEEE Computer Society Press, 54–64 (1995)

[Fagin 76] Fagin, R.: Probabilities on finite models, Probabilities on finite models, J. Symb. Logic, **41**, 50–58 (1976)

[Fagin 93] Fagin, R.: Finite-model theory—a personal perspective, Theoretical computer science, **116**, 3–31 (1993)

[Fernandez de la Vega 90] Fernandez de la Vega, W.: Kernels in Random Graphs, Discrete Mathematics, **82**, 213–217 (1990)

[Gaifman 64] Gaifman, H.: Concerning measures in first- order calculi, Israel Journal of Mathematics, **2**, 1–18 (1964)

[Glebskii et al 69] Glebskii, Y.V., Kogan, D.I., Liogonki, M.I., Talanov, V.A.: Range and degree of realizability of formulas in the restricted predicate calculus, Cybernetics, **5**, 142–154 (1969)

[Goranko and Passy 92] Goranko, V., Passy, S.: Using the universal modality: gains and questions, J. of Logic and Computation, **2**, 5–30 (1992)

[Grandjean 83] Grandjean, E.: Complexity of the first–order theory of almost all structures, Information and Computation, **52**, 180–204 (1983)

[Gurevich 92] Gurevich, Y.: Zero-One Laws: The Logic in Computer Science Column, Bulletin of the European Association for Theoretical Computer Science, 90–106 (1992)

[Halpern and Kapron 94] Halpern, J., Kapron, B.: Zero-one laws for modal logic, Ann. Pure and Appl. Logic, **69**, 157–193 (1994)

[Kapron 87] Kapron, B.: Modal sequents and definability, J. Symb. Logic, pp. **52**, 756–762 (1987)

[Knyazev 90] Knyazev, V.V.: Zero-one law for an extension of first-order predicate calculus, Kybernetika, **26**, 292–296 (1990)

[Kolaitis and Vardi 89] Kolaitis, Ph.G., Vardi, M.Y.: 0-1 Laws for Fragments of Second-Order Logic: an overview, In: Y. Moschovakis (ed.), *Logic for Computer Science*, Proc. of workshop in 1989, 265–286

[Kolaitis and Vardi 90] Kolaitis, Ph.G., Vardi, M.Y.: 0-1 Laws and Decision Problems for Fragments of Second-Order Logic, Information and Computation, **87**, 302–338 (1990)

[Kolaitis and Vardi 92] Kolaitis, Ph.G., Vardi, M.Y.: 0-1 laws for infinitary logics, Information and Computation, **98**, 250–294 (1992)

[Lacoste 96] Lacoste, T.: Finitistic proofs of 0-1 laws for fragments of second-order logic, Information Processing Letters, **58**, 1–4 (1996)

[Lynch 85] Lynch, J.: Probabilities of first-order sentences about unary functions, Trans. American Mathematical Society, **287**, 543–568 (1985)

[Palmer 85] Palmer, E.: *Graphical Evolution: An Introduction to the Theory of Random Graphs*, Wiley, 1985

[Spaan 93] Spaan, E.: Complexity of modal logics, Ph D. Thesis, ILLC, Amsterdam, 1993

[Talanov 81] Talanov, V.: Asymptotic Solvability of logical Formulas, in: *Combinatorial-Algebraic Methods in Applied Mathematics*, Gorky Univ., USSR, 1981 (in Russian), 118–126