ASYMPTOTIC CONDITIONAL PROBABILITIES: THE UNARY CASE*

ADAM J. GROVE[†], JOSEPH Y. HALPERN[‡], AND DAPHNE KOLLER[§]

Abstract. Motivated by problems that arise in computing *degrees of belief*, we consider the problem of computing asymptotic conditional probabilities for first-order sentences. Given first-order sentences φ and θ, we consider the structures with domain $\{1,\ldots,N\}$ that satisfy θ, and compute the fraction of them in which φ is true. We then consider what happens to this fraction as N gets large. This extends the work on 0-1 laws that considers the limiting probability of first-order sentences, by considering asymptotic *conditional* probabilities. As shown by Līogon'kiĭ [*Math. Notes Acad. USSR*, 6 (1969), pp. 856–861] and by Grove, Halpern, and Koller [*Res. Rep.* RJ 9564, IBM Almaden Research Center, San Jose, CA, 1993], in the general case, asymptotic conditional probabilities do not always exist, and most questions relating to this issue are highly undecidable. These results, however, all depend on the assumption that θ can use a nonunary predicate symbol. Līogon'kiĭ [*Math. Notes Acad. USSR*, 6 (1969), pp. 856–861] shows that if we condition on formulas θ involving unary predicate symbols only (but no equality or constant symbols), then the asymptotic conditional probability does exist and can be effectively computed. This is the case even if we place no corresponding restrictions on φ. We extend this result here to the case where θ involves equality and constants. We show that the complexity of computing the limit depends on various factors, such as the depth of quantifier nesting, or whether the vocabulary is finite or infinite. We completely characterize the complexity of the problem in the different cases, and show related results for the associated approximation problem.

Key words. asymptotic probability, 0-1 law, finite model theory, degree of belief, labeled structures, principle of indifference, complexity

AMS subject classifications. 03B48, 60C05, 68Q25

1. Introduction.

Suppose we have a sentence θ expressing facts that are known to be true, and another sentence φ whose truth is uncertain. Our knowledge θ is often insufficient to determine the truth of φ: both φ and its negation may be consistent with θ. In such cases, it can be useful to assign a *probability* to φ given θ. In a companion paper [23], we described our motivation for investigating this idea, and presented our general approach. We repeat some of this material below, to provide the setting for the results of this paper.

One important application of assigning probabilities to sentences—indeed, the one that has provided most of our motivation—is in the domain of decision theory and artificial intelligence. Consider an agent (or expert system) whose knowledge consists of some facts θ, who would like to assign a *degree of belief* to a particular statement φ. For example, a doctor may want to assign a degree of belief to the hypothesis that

*Received by the editors October 13, 1993; accepted for publication (in revised form) July 7, 1994. Some of this research was performed while Adam Grove and Daphne Koller were at Stanford University and at the IBM Almaden Research Center. A preliminary version of this paper appeared in *Proc. 24th ACM Symp. on Theory of Computing* [20]. This research was sponsored in part by an IBM Graduate Fellowship to Adam Grove, by a University of California President's Postdoctoral Fellowship to Daphne Koller, and by Air Force Office of Scientific Research contract F49620-91-C-0080. The United States Government is authorized to reproduce and distribute reprints for governmental purposes. Posted with permission from the SIAM Journal on Computing, Vol. 25, No. 1.

[†]NEC Research Institute, 4 Independence Way, Princeton, NJ 08540 (grove@research.nj.nec.com).

[‡]IBM Almaden Research Center, 650 Harry Rd., San Jose, CA 95120 (halpern@almaden.ibm.com).

[§]Computer Science Division, University of California at Berkeley, Berkeley, CA 94720 (daphne@cs.berkeley.edu).

a patient has a particular illness, based on the symptoms exhibited by the patient together with general information about symptoms and diseases. Since the actions the agent takes may depend crucially on this value, we would like techniques for computing degrees of belief in a principled manner.

The difficulty of defining a principled technique for computing the probability of φ given θ, and then actually computing that probability, depends in part on the language and logic being considered. In decision theory, applications often demand the ability to express statistical knowledge (for instance, correlations between symptoms and diseases) as well as first-order knowledge. Work in the field of 0-1 *laws* (which, as discussed below, is closely related to our own) has examined some higher-order logics as well as first-order logic. Nevertheless, the pure first-order case is still difficult, and is important because it provides a foundation for all extensions. In this paper and in [23] we address the problem of computing conditional probabilities in the first-order case. In a related paper [22], we consider the case of a first-order logic augmented with statistical knowledge.

The general problem of assigning probabilities to first-order sentences has been well studied (cf. [15] and [16]). In this paper, we investigate two specific formalisms for computing probabilities, based on the same basic approach. Our approach is itself an instance of a much older idea, known as *the principle of insufficient reason* [28] or *the principle of indifference* [26]. This states that all possibilities should be given equal probability, and was once regarded as one of the most basic principles of probability theory. (See [24] for a discussion of the history of the principle.) We use this idea to assign equal degrees of belief to all basic "situations" consistent with the known facts. The two formalisms we consider differ only in how they interpret "situation." We discuss this in more detail below.

In many applications, including the one of most interest to us, it makes sense to consider finite domains only. In fact, the case of most interest to us is the behavior of the formulas φ and θ over large finite domains. Similar questions also arise in the area of 0-1 *laws*. Our approach essentially generalizes the methods used in the work on 0-1 laws for first-order logic to the case of conditional probabilities. (See Compton's overview [8] for an introduction to this work.) Assume, without loss of generality, that the domain is $\{1, \ldots, N\}$ for some natural number N. As we said above, we consider two notions of "situation." In the *random-worlds method*, the possible situations are all the worlds, or first-order models, with domain $\{1, \ldots, N\}$ that satisfy the constraints θ. Based on the principle of indifference, we assume that all worlds are equally likely. To assign a probability to φ, we therefore simply compute the fraction of them in which the sentence φ is true. The random-worlds approach views each individual in $\{1, \ldots, N\}$ as having a distinct name (even though the name may not correspond to any constant in the vocabulary). Thus, two worlds that are isomorphic with respect to the symbols in the vocabulary are still treated as distinct situations. In some cases, however, we may believe that all relevant distinctions are captured by our vocabulary, and that isomorphic worlds are not truly distinct. The *random-structures method* attempts to capture this intuition by considering a situation to be a *structure*—an *isomorphism class* of worlds. This corresponds to assuming that individuals are distinguishable only if they differ with respect to properties definable by the language. As before, we assign a probability to φ by computing the fraction of the structures that satisfy φ among those structures that satisfy θ.[1]

[1] The random-worlds method considers what has been called in the literature *labeled* structures,

Since we are computing probabilities over finite models, we have assumed that the domain is $\{1, \ldots, N\}$ for some N. However, we often do not know the precise domain size N. In many cases, we know only that N is large. We therefore estimate the probability of φ given θ by the asymptotic limit, as N grows to infinity, of this probability over models of size N.

Precisely the same definitions of asymptotic probability are used in the context of 0-1 laws for first-order logic, but without allowing prior information θ. The original 0-1 law, proved independently by Glebskiĭ et al. [18] and Fagin [13], states that the asymptotic probability of any first-order sentence φ with no constant or function symbols is either 0 or 1. This means that such a sentence is true in almost all finite structures, or in almost none.

Our work differs from the original work on 0-1 laws in two respects. The first is relatively minor: we need to allow the use of constant symbols in φ, as they are necessary when discussing individuals (such as patients). Although this is a minor change, it is worth observing that it has a significant impact. It is easy to see that once we allow constant symbols, the asymptotic probability of a sentence φ is no longer either 0 or 1; for example, the asymptotic probability of $P(c)$ is $\frac{1}{2}$. Moreover, once we allow constant symbols, the asymptotic probability under random worlds and under random structures need not be the same. The more significant difference, however, is that we are interested in the asymptotic *conditional* probability of φ, given some prior knowledge θ. That is, we want the probability of φ over the class of finite structures defined by θ.

Some work has already been done on aspects of this question. Līogon'kīĭ [31], and independently Fagin [13], showed that asymptotic conditional probabilities do not necessarily converge to any limit. Subsequently, 0-1 laws were proved for special classes of first-order structures (such as graphs, tournaments, partial orders, etc.; see the overview paper [8] for details and further references). In many cases, the classes considered could be defined in terms of first-order constraints. Thus, these results can be viewed as special cases of the problem that we are interested in: computing asymptotic *conditional* probabilities relative to structures satisfying the constraints of a database. Lynch [32] showed that, for the random-worlds method, asymptotic probabilities exist for first-order sentences involving unary functions, although there is no 0-1 law. (Recall that the original 0-1 result is specifically for first-order logic *without* function symbols.) This can also be viewed as a special case of an asymptotic conditional probability for first-order logic without functions, since we can replace the unary functions by binary predicates, and condition on the fact that they are functions.

The most comprehensive work on this problem is the work of Līogon'kīĭ [31].[2] In addition to pointing out that asymptotic conditional probabilities do not exist in general, he shows that it is undecidable whether such a probability exists. (We generalize Līogon'kīĭ's results for this case in [23].) He then investigates the special case of conditioning on formulas involving unary predicates only (but no constants

while the random-structures method considers *unlabeled* structures [8]. We choose to use our own terminology for the random-worlds and random-structures methods, rather than the terminology of labeled and unlabeled. This is partly because we feel it is more descriptive, and partly because there are other variants of the approach that are useful for our intended application, and that do not fit into the standard labeled/unlabeled structures dichotomy (see [2]).

[2] In an earlier version of this paper [21], we stated that, to our knowledge, no work had been done on the general problem of asymptotic conditional probabilities. We thank Moshe Vardi for pointing out to us the work of Līogon'kīĭ [31].

or equality). In this case, he proves that the asymptotic conditional probability does exist and can be effectively computed, even if the left side of the conditional, φ, has predicates of arbitrary arity and equality. This gap between unary predicates and binary predicates is somewhat reminiscent of the fact that first-order logic over a vocabulary with only unary predicates (and constant symbols) is decidable, while if we allow even a single binary predicate symbol, then it becomes undecidable [11], [29]. This similarity is not coincidental; some of the techniques used to show that first-order logic over a vocabulary with unary predicate symbols is decidable are used by us to show that asymptotic conditional probabilities exist.

In this paper, we extend the results of Līogon'kīĭ [31] for the unary case. We first prove (in §3) that, if we condition on a formula involving only unary predicates, constants, and equality that is satisfiable in arbitrarily large models, the asymptotic conditional probability exists. We also present an algorithm for computing this limit. The main idea we use is the following: to compute the asymptotic conditional probability of φ given θ, we examine the behavior of φ in finite models of θ. It turns out that we can partition the models of θ into a finite collection of classes, such that φ behaves *uniformly* in any individual class. By this we mean that almost all worlds in the class satisfy φ or almost none do; i.e., there is a 0-1 law for the asymptotic probability of φ when we restrict attention to models in a single class. In §3 we define these classes and prove the existence of a 0-1 law within each class. We also show how the asymptotic conditional probability of φ given θ can be computed using these 0-1 probabilities.

In §4 we turn our attention to the complexity of computing the asymptotic probability in this case. Our results, which are the same for random worlds and random structures, depend on several factors: whether the vocabulary is finite or infinite, whether there is a bound on the depth of quantifier nesting, whether equality is used in θ, whether nonunary predicates are used, and whether there is a bound on predicate arities. For a fixed and finite vocabulary, there are just two cases: if there is no bound on the depth of quantifier nesting then computing asymptotic conditional probabilities is PSPACE-complete, otherwise the computation can be done in linear time. The case in which the vocabulary is not fixed (which is the case more typically considered in complexity theory) is more complex. The results for this case are summarized in Table 1.1. Perhaps the most interesting aspect of this table is the factors that cause the difference in complexity between #EXP and #TA(EXP,LIN) (where #TA(EXP,LIN) is the counting class corresponding to alternating Turing machines that take exponential time and make only a linear number of alternations; a formal definition is provided in §4.5). If we allow the use of equality in θ, then we need to restrict both φ and θ to using only unary predicates to get the #EXP upper bound. On the other hand, if θ does not mention equality, then the #EXP upper bound is attained as long as there is some fixed bound on the arity of the predicates appearing in φ. If we have no bound on the arity of the predicates that appear in φ, or if we allow equality in θ and predicates of arity 2 in φ, then the #EXP upper bound no longer holds, and we move to #TA(EXP,LIN).

Our results showing that computing the asymptotic probability is hard can be extended to show that finding a nontrivial estimate of the probability (i.e., deciding if it lies in a nontrivial interval) is almost as difficult. The lower bounds for the arity-bounded case and the general case require formulas of quantification depth 2 or more. For unquantified sentences or depth-1 quantification, things seem to become an exponential factor easier. We do not have tight bounds for the complexity of

TABLE 1.1
Complexity of asymptotic conditional probabilities.

	Depth ≤ 1	**Restricted**	**General case**
Existence	NP-complete	NEXPTIME-complete	NEXPTIME-complete
Compute	#P/PSPACE	#EXP-complete	#TA(EXP,LIN)-complete
Approximate	(co-)NP-hard	(co-)NEXPTIME-hard	TA(EXP,LIN)-hard

computing the degree of belief in this case; we have a #P lower bound and a PSPACE upper bound. The results for depth 1 are not proved in this paper; see [27] for details.

We observe that apart from our precise classification of the complexity of these problems, our analysis provides an effective algorithm for computing the asymptotic conditional probability. The complexity of this algorithm is, in general, double-exponential in the number of unary predicates used and in the maximum arity of any predicate symbol used; it is exponential in the overall size of the vocabulary and in the lengths of φ and θ.

Our results are of more than purely technical interest. The random-worlds method is of considerable theoretical and practical importance. We have already mentioned its relevance to computing degrees of belief. There are well-known results from physics that show the close connection between the random-worlds method and *maximum entropy* [25]. These results say that in certain cases the asymptotic probability can be computed using maximum entropy methods. Some formalization of similar results, but in a framework that is close to that of the current paper, can be found in [33] and [22]. (These results are of far more interest when there are statistical assertions in the language, so we do not discuss them here.)

As we observe in [23] and [22], this connection relies on the fact that we are conditioning on a unary formula. In fact, in almost all applications where maximum entropy has been used (and where its application can be best justified in terms of the random-worlds method) the knowledge base is described in terms of unary predicates (or, equivalently, unary functions with a finite range). For example, in physics applications we are interested in such predicates as quantum state (see [10]). Similarly, AI applications and expert systems typically use only unary predicates [7] such as symptoms and diseases. In general, many properties of interest can be expressed using unary predicates, since they express properties of individuals. Indeed, a good case can be made that statisticians tend to reformulate all problems in terms of unary predicates, since an event in a sample space can be identified with a unary predicate [36]. Indeed, in most cases where statistics are used, we have a basic unit in mind (an individual, a family, a household, etc.), and the properties (predicates) we consider are typically relative to a single unit (i.e., unary predicates). Thus, results concerning computing the asymptotic conditional probability if we condition on unary formulas are significant in practice.

2. Definitions. Let Φ be a set of predicate and function symbols, and let $\mathcal{L}(\Phi)$ (resp., $\mathcal{L}^-(\Phi)$) denote the set of first-order sentences over Φ with equality (resp., without equality). To simplify the presentation, we begin by assuming that Φ is finite; the case of an infinite vocabulary is deferred to §2.3. Much of the material in §§2.1 and 2.2 is taken from [23].

2.1. The random-worlds method. We begin by defining the random-worlds, or labeled, method. Given a sentence $\xi \in \mathcal{L}(\Phi)$, let $\#world_N^{\Phi}(\xi)$ be the number of worlds, or first-order models, of ξ over Φ with domain $\{1, \ldots, N\}$. Note that the assumption that Φ is finite is necessary for $\#world_N^{\Phi}(\xi)$ to be well defined. Define

$$\Pr_N^{w,\Phi}(\varphi \mid \theta) = \frac{\#world_N^{\Phi}(\varphi \wedge \theta)}{\#world_N^{\Phi}(\theta)} .$$

In [23], we proved the following proposition.

PROPOSITION 2.1. *Let Φ, Φ' be finite vocabularies, and let φ, θ be sentences in both $\mathcal{L}(\Phi)$ and $\mathcal{L}(\Phi')$. Then $\Pr_N^{w,\Phi}(\varphi \mid \theta) = \Pr_N^{w,\Phi'}(\varphi \mid \theta)$.*

Thus, the value of $\Pr_N^{w,\Phi}(\varphi \mid \theta)$ does not depend on the choice of Φ. We therefore omit reference to Φ in $\Pr_N^{w,\Phi}(\varphi \mid \theta)$, writing $\Pr_N^{w}(\varphi \mid \theta)$ instead.

We would like to define $\Pr_\infty^{w}(\varphi \mid \theta)$ as the limit $\lim_{N\to\infty} \Pr_N^{w}(\varphi \mid \theta)$. There is a small technical problem we have to deal with in this definition: we must decide what to do if $\#world_N^{\Phi}(\theta) = 0$, so that $\Pr_N^{w}(\varphi \mid \theta)$ is not well defined. In [23], we differentiate between the case where $\Pr_N^{w}(\varphi \mid \theta)$ is well defined for all but finitely many N's, and the case where it is well defined for infinitely many N's. As we shall show (see Lemma 3.30) this distinction need not be made when θ is a unary formula. Thus, for the purposes of this paper, we use the following definition of well-definedness, which is simpler than that of [23].

DEFINITION 2.2. *The asymptotic conditional probability according to the random-worlds method, denoted $\Pr_\infty^{w}(\varphi \mid \theta)$, is* well defined *if $\#world_N^{\Phi}(\theta) \neq 0$ for all but finitely many N. If $\Pr_\infty^{w}(\varphi \mid \theta)$ is well defined, then we take $\Pr_\infty^{w}(\varphi \mid \theta)$ to denote $\lim_{N\to\infty} \Pr_N^{w}(\varphi \mid \theta)$.* □

Note that for any formula φ, the issue of whether $\Pr_\infty^{w}(\varphi \mid \theta)$ is well defined is completely determined by θ. Therefore, when investigating the question of how to decide whether such a probability is well defined, it is often useful to ignore φ. We therefore say that $\Pr_\infty^{w}(* \mid \theta)$ *is well defined* if $\Pr_\infty^{w}(\varphi \mid \theta)$ is well defined for every formula φ (which is true iff $\Pr_\infty^{w}(\mathit{true} \mid \theta)$ is well defined).

2.2. The random-structures method. As we explained in the introduction, the random-structures method is motivated by the intuition that worlds that are indistinguishable within the language should only be counted once. Thus, the random-structures method counts the number of (*unlabeled*) *structures*, or isomorphism classes of worlds.

Formally, we proceed as follows. Given a sentence $\xi \in \mathcal{L}(\Phi)$, let $\#struct_N^{\Phi}(\xi)$ be the number of isomorphism classes of worlds with domain $\{1, \ldots, N\}$ over the vocabulary Φ satisfying ξ. Note that since all the worlds that make up a structure agree on the truth value they assign to ξ, it makes sense to talk about a structure satisfying or not satisfying ξ. We can then proceed, as before, to define $\Pr_N^{s,\Phi}(\varphi \mid \theta)$ as $\frac{\#struct_N^{\Phi}(\varphi\wedge\theta)}{\#struct_N^{\Phi}(\theta)}$. We define asymptotic conditional probability, denoted $\Pr_\infty^{s,\Phi}(\varphi \mid \theta)$ in terms of $\Pr_N^{s,\Phi}(\varphi \mid \theta)$, in analogy to the earlier definition for random worlds. It is clear that $\#world_N^{\Phi}(\theta) = 0$ iff $\#struct_N^{\Phi}(\theta) = 0$, so that well-definedness is equivalent for the two methods, for any φ, θ.

PROPOSITION 2.3. *For any $\theta \in \mathcal{L}(\Phi)$, $\Pr_\infty^{w}(* \mid \theta)$ is well defined iff $\Pr_\infty^{s,\Phi}(* \mid \theta)$ is well defined.*

As the following example, taken from [23], shows, for the random-structures method the analogue to Proposition 2.1 does not hold; the value of $\Pr_N^{s,\Phi}(\varphi \mid \theta)$, and even the value of the limit, depends on the choice of Φ. This example, together with Proposition 2.1, also demonstrates that the values of conditional probabilities generally differ between the random-worlds method and the random-structures method.

By way of contrast, Fagin [14] showed that the random-worlds and random-structures methods give the same answers for unconditional probabilities, if we do not have constant or function symbols in the language.

Example 2.4. Suppose $\Phi = \{P\}$, where P is a unary predicate symbol. Let θ be $\exists! x\, P(x) \vee \neg\exists x\, P(x)$ (where, as usual, "$\exists!$" means "exists a unique"), and let φ be $\exists x\, P(x)$. For any domain size N, $\#struct_N^{\Phi}(\theta) = 2$. In one structure, there is exactly one element satisfying P and $N-1$ satisfying $\neg P$; in the other, all elements satisfy $\neg P$. Therefore, $\Pr_\infty^{s,\Phi}(\varphi \mid \theta) = \frac{1}{2}$.

Now, consider $\Phi' = \{P, Q\}$, for a new unary predicate Q. There are $2N$ structures where there exists an element satisfying P: the element satisfying P may or may not satisfy Q, and of the $N-1$ elements satisfying $\neg P$, any number between 0 and $N-1$ may also satisfy Q. On the other hand, there are $N+1$ structures where all elements satisfy $\neg P$: any number of elements between 0 and N may satisfy Q. Therefore, $\Pr_N^{s,\Phi'}(\varphi \mid \theta) = \frac{2N}{3N+1}$, and $\Pr_\infty^{s,\Phi'}(\varphi \mid \theta) = \frac{2}{3}$.

We know that the asymptotic limit for the random-worlds method will be the same, whether we use Φ or Φ'. Using Φ, notice that the single structure where $\exists! x\, P(x)$ is true contains N worlds (corresponding to the choice of element satisfying P), whereas the other possible structure contains only one world. Therefore, $\Pr_\infty^{w}(\varphi \mid \theta) = 1$. $\square$

Although the two methods give different answers in general, we shall see in the next section that there are important circumstances under which they agree.

2.3. Infinite vocabularies. Up to now we have assumed that the vocabulary Φ is finite. As we observed, this assumption is crucial in our definitions of $\#world_N^{\Phi}(\xi)$ and $\#struct_N^{\Phi}(\xi)$. Nevertheless, in many standard complexity arguments it is important that the vocabulary be infinite. For example, satisfiability for propositional logic formulas is decidable in linear time if we consider a single finite vocabulary; we need to consider the class of formulas definable over some infinite vocabulary of propositional symbols to get NP-completeness.

How can we modify the random-worlds and random-structures methods to deal with an infinite vocabulary Ω? The issue is surprisingly subtle. One plausible choice depends on the observation that even if Ω is infinite, the set of symbols appearing in a given sentence is always finite. We can thus do our computations relative to this set. More formally, if $\Omega_{\varphi\wedge\theta}$ denotes the set of symbols in Ω that actually appear in $\varphi \wedge \theta$, we could define $\Pr_N^{w,\Omega}(\varphi \mid \theta) = \Pr_N^{w,\Omega_{\varphi\wedge\theta}}(\varphi \mid \theta)$. Similarly, for the random-structures method, we could define $\Pr_N^{s,\Omega}(\varphi \mid \theta) = \Pr_N^{s,\Omega_{\varphi\wedge\theta}}(\varphi \mid \theta)$. The problem with this approach is that the values given by the random-structures approach depend on the vocabulary, and it is easy to find two equivalent sentences φ and φ' such that $\Omega_\varphi \neq \Omega_{\varphi'}$ and $\Pr_\infty^{s,\Omega_{\varphi\wedge\theta}}(\varphi \mid \theta) \neq \Pr_\infty^{s,\Omega_{\varphi'\wedge\theta}}(\varphi' \mid \theta)$. (A simple example of this phenomenon can be obtained by modifying Example 2.4 slightly.) Thus, under this approach, the value of asymptotic conditional probabilities can depend on the precise syntax of the sentences involved. We view this as undesirable, and so we focus on the following two interpretations of infinite vocabularies.[3]

The first of these two alternative approaches treats an infinite vocabulary as a limit of finite subvocabularies. Assume for ease of exposition that Ω is countable. Let Ω_m consist of the first m symbols in Ω (using some arbitrary fixed ordering). We can then define $\Pr_N^{w,\Omega}(\varphi \mid \theta) = \lim_{m\to\infty} \Pr_N^{w,\Omega_m}(\varphi \mid \theta)$ (where we

[3]We note, however, that all our later complexity results concerning infinite vocabularies can be easily shown to hold for the definition just discussed.

take $\Pr_N^{w,\Omega_m}(\varphi \mid \theta)$ to be undefined if $\varphi, \theta \notin \mathcal{L}(\Omega_m)$).[4] Similarly, we can define $\Pr_N^{s,\Omega}(\varphi \mid \theta) = \lim_{m\to\infty} \Pr_N^{s,\Omega_m}(\varphi \mid \theta)$. It follows from the results we prove below that these limits are independent of the ordering of the symbols in the vocabulary.

The second interpretation is quite different. In it, although there may be an infinite vocabulary Ω in the background, we assume that each problem instance comes along with a finite vocabulary Φ as part of the input. Thus, in our infinite vocabulary Ω, we may have predicates that are relevant to medical diagnosis, physics experiments, automobile insurance, etc. When thinking about medical applications, we use that finite portion Φ of the infinite vocabulary that is appropriate. In this approach, we always deal with finite vocabularies, but ones whose size is potentially unbounded because we do not fix the relevant vocabulary in advance.

In essence, the first approach can be viewed as saying that there really is an infinite vocabulary, while the second approach considers there to be an infinite collection of finite vocabularies (with no bound on the size of the vocabularies in the collection). The distinction between these possibilities is not usually examined as closely as we have done here. This is because the difference is rarely important. For example, propositional satisfiability is NP-complete over an infinite vocabulary, no matter how we interpret "infinite." In our context, the difference turns out to be moderately significant. For random worlds, an argument based on Proposition 2.1 shows the two approaches lead to the same answers (as does the approach that we rejected where, when computing $\Pr_N^{w,\Omega}(\varphi \mid \theta)$, we restrict the vocabulary to $\Omega_{\varphi\wedge\theta}$). On the other hand, the two approaches can lead to quite different answers in the case of the random-structures approach. It is important to point out, however, that the complexity of all problems we consider turns out to be the same no matter which interpretation of "infinite" we use.

In fact, as we now show, according to the first approach the random-structures method and the random-worlds method agree whenever we have an infinite vocabulary (and thus we have an analogue to Fagin's result [14] for the case of unconditional probabilities). A structure of size N is an equivalence class of at most $N!$ worlds, since there are at most $N!$ worlds isomorphic to a given world. We say that such a structure is *rigid* if it consists of exactly $N!$ worlds. It is easy to see that a structure is rigid just if every (nontrivial) permutation of the domain elements in a world that makes up the structure produces a *different* world in that structure. We say a world is rigid if the corresponding structure is.

Example 2.5. Let Φ consist of a single unary predicate P, and consider the worlds over the domain $\{1, 2, 3\}$. All worlds where the denotation of P contains exactly two elements are isomorphic. Therefore, these worlds form a single structure $\mathcal{S}$. There are three worlds in $\mathcal{S}$, corresponding to the possible denotations of P: $\{1, 2\}$, $\{1, 3\}$, $\{2, 3\}$. Therefore, $\mathcal{S}$ is not rigid. In fact, it is easy to see that no structure over Φ is rigid. Now, consider structures over $\Phi' = \{P, Q\}$, where Q is a new unary predicate. The set of all worlds where the denotation of P contains two elements no longer forms a structure over Φ'. For example, one structure $\mathcal{S}'$ over Φ' is the set of worlds where the denotations of $P \wedge Q$, $P \wedge \neg Q$, and $\neg P \wedge Q$ each contain one element. There are six worlds in $\mathcal{S}'$, corresponding to the possible permutations of the three domain elements. Therefore, $\mathcal{S}'$ is rigid. $\square$

This example demonstrates that increasing the vocabulary tends to cause rigidity.

[4]Here, we chose to take the limit on the vocabulary, and only then to take the limit on the domain size. We could, however, have chosen to reverse the order of the limits, or to consider arbitrary joint limits of these two parameters. The approach taken here seems to be the most well motivated in this framework.

We now formalize this intuition, and show its importance. Note that in the following definition (and throughout the paper) all logarithms are taken to the base 2.

DEFINITION 2.6. *We say that a vocabulary* Φ *is* sufficiently rich *with respect to* N *if*

(a) Φ *contains at least* κ_N *constant symbols and* $\kappa_N \geq N^2 \log N$*, or*

(b) Φ *contains at least* π_N *unary predicate symbols and* $\pi_N \geq 3 \log N$*, or*

(c) Φ *contains at least one nonunary predicate symbol.* □

Fagin showed that if Φ contains at least one nonunary predicate symbol, then the number of worlds over Φ of size N is asymptotically $N!$ times the number of structures [14]. That is, almost all structures are rigid in this case. We now generalize this result. Let *rigid* be an assertion that is true only in rigid structures or rigid worlds; note that *rigid* cannot be expressed in first-order logic. If $F(N)$ and $G(N)$ are two functions of N, we write $F(N) \sim G(N)$ if $\lim_{N\to\infty} F(N)/G(N) = 1$.

THEOREM 2.7. *Suppose that for every* N*,* Φ *and* Ω_N *are disjoint finite vocabularies such that* Ω_N *is sufficiently rich with respect to* N*. Then for any* $\xi \in \mathcal{L}(\Phi)$*,*

$$\lim_{N\to\infty} \Pr_N^{s,\Phi\cup\Omega_N}(\mathit{rigid} \mid \xi) = 1,$$

provided that ξ *is satisfiable for all sufficiently large domains. Hence,* $\#\mathit{world}_N^{\Phi\cup\Omega_N}(\xi) \sim N!\#\mathit{struct}_N^{\Phi\cup\Omega_N}(\xi)$*.*

Proof. We first prove the result under the additional assumptions that $\xi = \mathit{true}$ and $\Phi = \emptyset$. We consider each of the three possibilities for sufficient richness separately, and for each case we show that almost all structures are rigid. As we said above, the case where Ω_N contains at least one nonunary predicate and $\xi = \mathit{true}$ is Fagin's result, so we need only consider the remaining two cases.

Suppose $\xi = \mathit{true}$, $\Phi = \emptyset$, and Ω_N contains κ_N constant symbols. Without loss of generality, we can assume that these constants are the only symbols in Ω_N, because any expansion of a rigid structure over Ω_N to a richer vocabulary will also be rigid. Consider a structure $\mathcal{S}$. All the worlds that make up $\mathcal{S}$ must agree on the equality relations between the interpretations of the constants. That is, for any pair of constant symbols c and c', either they are equal in all worlds that make up the structure or not equal in all of them. Thus, a lower bound on the number of distinct structures over Ω_N is given by the number of ways of partitioning κ_N objects into N or fewer equivalence classes. There is no closed form expression for this number, but a simple lower bound is obtained by counting structures where the first N constants denote distinct objects. There are $N^{(\kappa_N - N)}$ such structures, because we must choose, for each of the other constants, to which of the first N constants it is equal. It is easy to see that if all or all but one of the elements in a structure (that is, in any of the worlds in that structure) are denoted by some constant, then this structure is rigid. Hence, if a structure is nonrigid, then two or more elements are not denoted by any constant. Thus, an upper bound on the number of nonrigid structures is $(N-2)^{\kappa_N}$. Therefore,

$$\Pr_N^{s,\Omega_N}(\neg\mathit{rigid} \mid \mathit{true}) \leq \frac{(N-2)^{\kappa_N}}{N^{\kappa_N - N}} = N^N \left(1 - \frac{2}{N}\right)^{\kappa_N} < N^N e^{-2\frac{\kappa_N}{N}}.$$

This will tend to 0 if $\kappa_N \geq N^2 \log N$.

Next, suppose that $\xi = \mathit{true}$, $\Phi = \emptyset$, and Ω_N contains π_N unary predicate symbols. As before, we can assume that these predicates comprise all of Ω_N. Consider a structure $\mathcal{S}$ and a world $\mathcal{W}$ in the isomorphism class making up that structure. These π_N unary predicates partition the domain of $\mathcal{W}$ into 2^{π_N} cells, according to the subset of predicates satisfied by each of the domain elements. Notice that the predicates actually partition each of the isomorphic worlds in s in the same way (in that corresponding elements of the partition have the same size). Thus, a lower bound on the number of distinct structures over Φ is the number of ways of allocating N indistinguishable elements into 2^{π_N} distinguishable cells, which is $\binom{2^{\pi_N}+N-1}{N}$. Clearly, a structure is nonrigid if and only if some element of the partition contains more than one domain element. Thus, an upper bound on the number of nonrigid structures can be obtained by counting the number of structures over $N-1$ elements, then choosing one of the these elements to be a "double" element, representing two elements. This can be done in $(N-1)\binom{2^{\pi_N}+N-2}{N-1}$ ways. Therefore,

$$\Pr_N^{s,\Omega_N}(\neg \mathit{rigid} \mid \mathit{true}) \leq \frac{(N-1)\binom{2^{\pi_N}+N-2}{N-1}}{\binom{2^{\pi_N}+N-1}{N}} = \frac{N^2-N}{2^{\pi_N}+N-1}.$$

This tends to zero if $2^{\pi_N}/N^2 \to \infty$ as N grows, which is ensured by the assumption $\pi_N \geq 3\log N$.

Finally, we drop the assumptions that $\xi = \mathit{true}$ and $\Phi = \emptyset$. Given a structure over Ω_N, we can choose the denotation for the predicates in Φ in any way that satisfies ξ. It is easy to verify that if the original structure is rigid, all such choices lead to distinct structures. Therefore,

$$\#\mathit{struct}_N^{\Phi\cup\Omega_N}(\mathit{rigid} \wedge \xi) \geq \#\mathit{struct}_N^{\Omega_N}(\mathit{rigid}) \cdot \#\mathit{world}_N^{\Phi}(\xi)\ .$$

On the other hand, clearly

$$\#\mathit{struct}_N^{\Phi\cup\Omega_N}(\neg\mathit{rigid} \wedge \xi) \leq \#\mathit{struct}_N^{\Omega_N}(\neg\mathit{rigid}) \cdot \#\mathit{world}_N^{\Phi}(\xi)\ .$$

The second factor is the same in both these bounds, and therefore

$$\Pr_N^{s,\Phi\cup\Omega_N}(\mathit{rigid} \mid \xi) \geq \Pr_N^{s,\Omega_N}(\mathit{rigid} \mid \mathit{true})\ .$$

From our results for $\xi = \mathit{true}$ and $\Phi = \emptyset$ we conclude that

$$\lim_{N\to\infty} \Pr_N^{s,\Phi\cup\Omega_N}(\mathit{rigid} \mid \xi) = 1.$$

□

We also need to prove an analogous result for the random-worlds method. Note that while, if we restrict to formulas in $\mathcal{L}(\Phi)$, the answers given by the random-worlds method are independent of the vocabulary, the predicate *rigid* has a special definition in terms of the random-structures method, and so rigidity may well depend on the vocabulary. Thus, in the next result, we are careful to mention the vocabulary being used.

COROLLARY 2.8. *Suppose that for every N, Φ and Ω_N are disjoint finite vocabularies such that Ω_N is sufficiently rich with respect to N. Then for any $\xi \in \mathcal{L}(\Phi)$,*

$$\lim_{N\to\infty} \Pr_N^{w,\Phi\cup\Omega_N}(\mathit{rigid} \mid \xi) = 1,$$

provided that ξ is satisfiable in all sufficiently large domains.

Proof. Any rigid structure with domain size N that satisfies ξ corresponds to $N!$ worlds. On the other hand, nonrigid structures correspond to fewer than $N!$ worlds. It follows that the proportion of worlds satisfying ξ that are rigid is at least as great as the proportion of structures satisfying ξ that are rigid. Since the latter proportion is asymptotically 1, so is the former. □

Our main use of this theorem is in the following two corollaries. The first shows that when the vocabulary is infinite (and therefore sufficiently rich) the random-worlds and random-structures methods coincide. The second corollary shows that the same thing happens when the vocabulary is sufficiently rich because of a high-arity predicate, as long as this predicate does not appear in the formula we are conditioning on.

COROLLARY 2.9. *Suppose that* Ω *is infinite and* $\varphi, \theta \in \mathcal{L}(\Omega)$. *Then*

(a) $\Pr_N^{w,\Omega}(\varphi \mid \theta) \sim \Pr_N^{s,\Omega}(\varphi \mid \theta)$,

(b) $\Pr_\infty^{w,\Omega}(\varphi \mid \theta) = \Pr_\infty^{s,\Omega}(\varphi \mid \theta)$.

Proof. Fix N, and let Ω_m be the first m symbols in some enumeration of Ω. We will be interested in the limit as $m \longrightarrow \infty$, so without loss of generality assume that $m > N^2 \log N + |\Omega_{\varphi\wedge\theta}|$. Clearly $\Omega_m - \Omega_{\varphi\wedge\theta}$ is sufficiently rich with respect to N, so by Theorem 2.7, almost all structures are rigid. Since a rigid structure over a domain of size N consists of $N!$ worlds, we get:

$$\Pr_N^{w,\Omega_{\varphi\wedge\theta}\cup\Omega_m}(\varphi \mid \theta) = \frac{\#world_N^{\Omega_{\varphi\wedge\theta}\cup\Omega_m}(\varphi\wedge\theta)}{\#world_N^{\Omega_{\varphi\wedge\theta}\cup\Omega_m}(\theta)} \sim \frac{\#struct_N^{\Omega_{\varphi\wedge\theta}\cup\Omega_m}(\varphi\wedge\theta)}{\#struct_N^{\Omega_{\varphi\wedge\theta}\cup\Omega_m}(\theta)}$$

$$= \Pr_N^{s,\Omega_{\varphi\wedge\theta}\cup\Omega_m}(\varphi \mid \theta).$$

Since this holds for any sufficiently large m, it certainly holds at the limit. This proves part (a). Part (b) follows easily. □

We can easily strengthen part (a) and prove that we actually have $\Pr_N^{w,\Omega}(\varphi \mid \theta) = \Pr_N^{s,\Omega}(\varphi \mid \theta)$, for all N. Since we do not need this result in this paper, we omit the proof here. We remark that this result also holds for much richer languages; we did not use the fact that we were dealing with first-order logic anywhere in the proof.

COROLLARY 2.10. *Suppose that* $\varphi, \theta \in \mathcal{L}(\Phi)$ *where* Φ *contains some nonunary predicate symbol that does not appear in* θ. *Then* $\Pr_\infty^{w}(\varphi \mid \theta) = \Pr_\infty^{s,\Phi}(\varphi \mid \theta)$.

Proof. Using the rules of probability theory, we know that

$$\Pr_\infty^{s,\Phi}(\varphi \mid \theta) = \Pr_\infty^{s,\Phi}(\varphi \mid \theta\wedge rigid)\cdot\Pr_\infty^{s,\Phi}(rigid \mid \theta) + \Pr_\infty^{s,\Phi}(\varphi \mid \theta\wedge\neg rigid)\cdot\Pr_\infty^{s,\Phi}(\neg rigid \mid \theta)\ ,$$

if all limits exist. Because of the high-arity predicate, $\Phi - \Phi_\theta$ is sufficiently rich with respect to any N. Therefore, by Theorem 2.7, we deduce that $\Pr_\infty^{s,\Phi}(rigid \mid \theta) = 1$ and $\Pr_\infty^{s,\Phi}(\neg rigid \mid \theta) = 0$. Thus

$$\Pr_\infty^{s,\Phi}(\varphi \mid \theta) = \Pr_\infty^{s,\Phi}(\varphi \mid \theta \wedge rigid)\ .$$

Using Corollary 2.8 instead of Theorem 2.7, we can similarly show

$$\Pr_\infty^{w,\Phi}(\varphi \mid \theta) = \Pr_\infty^{w,\Phi}(\varphi \mid \theta \wedge rigid)\ .$$

Because of rigidity,

$$\Pr_\infty^{s,\Phi}(\varphi \mid \theta \wedge rigid) = \Pr_\infty^{w,\Phi}(\varphi \mid \theta \wedge rigid).$$

The result now follows immediately. □

3. Asymptotic probabilities. We begin by defining some notation that will be used consistently throughout the rest of the paper. We use Φ to denote a finite vocabulary, which may include nonunary as well as unary predicate symbols and constant symbols. We take $\mathcal{P}$ to be the set of all unary predicates in Φ, $\mathcal{C}$ to be the set of all constant symbols in Φ, and define $\Psi = \mathcal{P} \cup \mathcal{C}$. Finally, if φ is a formula, we use Φ_φ to denote those symbols in Φ that appear in φ; we can similarly define $\mathcal{C}_\varphi$, $\mathcal{P}_\varphi$, and Ψ_φ.

Our goal is to show how to compute asymptotic conditional probabilities. As we explained in the introduction, the main idea is the following. To compute $\Pr^w_\infty(\varphi \mid \theta)$, we partition the models of θ into a finite collection of classes, such that φ behaves *uniformly* in any individual class, that is, there is a 0-1 law for the asymptotic probability of φ when we restrict attention to models in a single class. Computing $\Pr^w_\infty(\varphi \mid \theta)$ reduces to first identifying the classes, computing the relative weight of each class (which is required because the classes are not necessarily of equal relative size), and then deciding, for each class, whether the asymptotic probability of φ is zero or one. In this section we deal with the logical aspects of this process; namely, showing how to construct an appropriate partition into classes. In the next section, we use results from this section to construct algorithms that compute asymptotic probabilities, and examine the complexity of these algorithms.

For most of this section, we will concentrate on the asymptotic probability according to random worlds. In §3.5 we discuss the modifications needed to deal with random structures, which are relatively minor.

3.1. Unary vocabularies and atomic descriptions. The success of the approach outlined above depends on the lack of expressivity of unary languages. In this section we show that sentences in $\mathcal{L}(\Psi)$ can only assert a fairly limited class of constraints. For instance, one corollary of our general result will be the well-known theorem that, if $\theta \in \mathcal{L}(\Psi)$ is satisfiable at all, it is satisfiable in a "small" model, one of size at most exponential in the size of the θ. (See [1] for a proof of this result and further historical references.)

We start with some definitions.

DEFINITION 3.1. *Given a vocabulary Φ and a finite set of variables $\mathcal{X}$, a* complete description *D over Φ and $\mathcal{X}$ is an unquantified conjunction of formulas such that*

- *for every predicate $R \in \Phi \cup \{=\}$ of arity m, and for every $z_1, \ldots, z_m \in \mathcal{C} \cup \mathcal{X}$, D contains exactly one of $R(z_1, \ldots, z_m)$ or $\neg R(z_1, \ldots, z_m)$ as a conjunct;*
- *D is consistent.*[5] □

We can think of a complete description as being a formula that describes as fully as possible the behavior of the predicate symbols in Φ over the constant symbols in Φ and the variables in $\mathcal{X}$.

We can also consider complete descriptions over subsets of Φ. The case when we look just at the unary predicates and a single variable x will be extremely important.

DEFINITION 3.2. *Let $\mathcal{P}$ be $\{P_1, \ldots, P_k\}$. An* atom *over $\mathcal{P}$ is a complete description over $\mathcal{P}$ and some variable $\{x\}$. More precisely, it is a conjunction of the form $P_1'(x) \wedge \ldots \wedge P_k'(x)$, where each P_i' is either P_i or $\neg P_i$. Since the variable x is irrelevant to our concerns, we typically suppress it and describe an atom as a conjunction of the form $P_1' \wedge \ldots \wedge P_k'$.* □

Note that there are $2^k = 2^{|\mathcal{P}|}$ atoms over $\mathcal{P}$, and that they are mutually exclusive

[5] Inconsistency is possible because of the use of equality. For example, if D includes $z_1 = z_2$ as well as both $R(z_1, z_3)$ and $\neg R(z_2, z_3)$, it is inconsistent.

and exhaustive. We use $A_1, \dots, A_{2^{|\mathcal{P}|}}$ to denote the atoms over $\mathcal{P}$, listed in some fixed order. For example, there are four atoms over $\mathcal{P} = \{P_1, P_2\}$: $A_1 = P_1 \wedge P_2$, $A_2 = P_1 \wedge \neg P_2$, $A_3 = \neg P_1 \wedge P_2$, $A_4 = \neg P_1 \wedge \neg P_2$.

We now want to define the notion of *atomic description* which is, roughly speaking, a maximally expressive formula in the unary vocabulary Ψ. Fix a natural number M. A size M atomic description consists of two parts. The first part, the *size description with bound M*, specifies exactly how many elements in the domain should satisfy each atom A_i, except that if there are M or more elements satisfying the atom it only expresses that fact, rather than giving the exact count. More formally, given a formula $\xi(x)$ with a free variable x, we take $\exists^m x\, \xi(x)$ to be the sentence that says there are precisely m domain elements satisfying ξ:

$$\exists^m x\, \xi(x) =_{\mathrm{def}} \exists x_1 \, \dots \, x_m \left(\bigwedge_i \left(\xi(x_i) \wedge \bigwedge_{j \neq i} (x_j \neq x_i) \right) \wedge \forall y (\xi(y) \Rightarrow \vee_i (y = x_i)) \right).$$

Similarly, we define $\exists^{\geq m} x\, \xi(x)$ to be the formula that says that there are at least m domain elements satisfying ξ:

$$\exists^{\geq m} x\, \xi(x) =_{\mathrm{def}} \exists x_1 \, \dots \, x_m \left(\bigwedge_i \left(\xi(x_i) \wedge \bigwedge_{j \neq i} (x_j \neq x_i) \right) \right).$$

DEFINITION 3.3. *A* size description with bound M (over $\mathcal{P}$) *is a conjunction of $2^{|\mathcal{P}|}$ formulas: for each atom A_i over $\mathcal{P}$, it includes either $\exists^{\geq M} x\, A_i(x)$ or a formula of the form $\exists^m x\, A_i(x)$ for some $m < M$.* □

The second part of an atomic description is a complete description that specifies the properties of constants and free variables.

DEFINITION 3.4. *A* size M atomic description *(over Ψ and $\mathcal{X}$) is a conjunction of:*

- *a size description with bound M over $\mathcal{P}$, and*
- *a complete description over Ψ and $\mathcal{X}$.* □

Note that an atomic description is a finite formula, and there are only finitely many size M atomic descriptions over Ψ and $\mathcal{X}$ (for fixed M). For the purposes of counting atomic descriptions (as we do in §3.4), we assume some arbitrary but fixed ordering of the conjuncts in an atomic description. Under this assumption, we cannot have two distinct atomic descriptions that differ only in the ordering of conjuncts. Given this, it is easy to see that atomic descriptions are mutually exclusive. Moreover, atomic descriptions are exhaustive—the disjunction of all consistent atomic descriptions of size M is valid.

Example 3.5. Consider the following size description σ with bound 4 over $\mathcal{P} = \{P_1, P_2\}$:

$$\exists^1 x\, A_1(x) \wedge \exists^3 x\, A_2(x) \wedge \exists^{\geq 4} x\, A_3(x) \wedge \exists^{\geq 4} x\, A_4(x).$$

Let $\Psi = \{P_1, P_2, c_1, c_2, c_3\}$. It is possible to augment σ into an atomic description in many ways. For example, one consistent atomic description ψ_* of size 4 over Ψ and $\emptyset$ (no free variables) is:[6]

$$\sigma \wedge A_2(c_1) \wedge A_3(c_2) \wedge A_3(c_3) \wedge c_1 \neq c_2 \wedge c_1 \neq c_3 \wedge c_2 = c_3.$$

[6]In our examples, we use the commutativity of equality in order to avoid writing down certain superfluous disjuncts. In this example, for instance, we do not write down both $c_1 \neq c_2$ and $c_2 \neq c_1$.

On the other hand, the atomic description

$$\sigma \wedge A_1(c_1) \wedge A_1(c_2) \wedge A_3(c_3) \wedge c_1 \neq c_2 \wedge c_1 \neq c_3 \wedge c_2 \neq c_3$$

is an inconsistent atomic description, since σ dictates that there is precisely one element in the atom A_1, whereas the second part of the atomic description implies that there are two distinct domain elements in that atom. □

As we explained, an atomic description is, intuitively, a maximally descriptive sentence over a unary vocabulary. The following theorem formalizes this idea by showing that each unary formula is equivalent to a disjunction of atomic descriptions. For a given M and set $\mathcal{X}$ of variables, let $\mathcal{A}^{\Psi}_{M,\mathcal{X}}$ be the set of consistent atomic descriptions of size M over Ψ and $\mathcal{X}$.

DEFINITION 3.6. *Let $d(\xi)$ denote the* depth of quantifier nesting *in ξ. We define $d(\xi)$ by induction on the structure of ξ as follows:*

- $d(\xi) = 0$ *for any atomic formula ξ,*
- $d(\neg\xi) = d(\xi)$,
- $d(\xi_1 \wedge \xi_2) = d(\xi_1 \vee \xi_2) = \max(d(\xi_1), d(\xi_2))$,
- $d(\forall y\, \xi) = d(\exists y\, \xi) = d(\xi) + 1$. □

THEOREM 3.7. *If ξ is a formula in $\mathcal{L}(\Psi)$ whose free variables are contained in $\mathcal{X}$, and $M \geq d(\xi) + |\mathcal{C}| + |\mathcal{X}|$, then there exists a set of atomic descriptions $\mathcal{A}^{\Psi}_{\xi} \subseteq \mathcal{A}^{\Psi}_{M,\mathcal{X}}$ such that ξ is equivalent to $\bigvee_{\psi \in \mathcal{A}^{\Psi}_{\xi}} \psi$.*

Proof. We proceed by a straightforward induction on the structure of ξ. We assume without loss of generality that ξ is constructed from atomic formulas using only the operators $\wedge$, $\neg$, and $\exists$.

First suppose that ξ is an atomic formula. That is, ξ is either of the form $P(z)$ or of the form $z = z'$, for $z, z' \in \mathcal{C} \cup \mathcal{X}$. In this case, either the formula ξ or its negation appears as a conjunct in each atomic description $\psi \in \mathcal{A}^{\Psi}_{M,\xi}$. Let $\mathcal{A}^{\Psi}_{\xi}$ be those atomic descriptions in which ξ appears as a conjunct. Clearly, ξ is inconsistent with the remaining atomic descriptions. Since the disjunction of the atomic descriptions in $\mathcal{A}^{\Psi}_{M,\mathcal{X}}$ is valid, we obtain that ξ is equivalent to $\bigvee_{\psi \in \mathcal{A}^{\Psi}_{\xi}} \psi$.

If ξ is of the form $\xi_1 \wedge \xi_2$, then by the induction hypothesis, ξ_i is equivalent to the disjunction of a set $\mathcal{A}^{\Psi}_{\xi_i} \subseteq \mathcal{A}^{\Psi}_{M,\mathcal{X}}$, for $i = 1, 2$. Clearly ξ is equivalent to the disjunction of the atomic descriptions in $\mathcal{A}^{\Psi}_{\xi_1} \cap \mathcal{A}^{\Psi}_{\xi_2}$. (Recall that the empty disjunction is equivalent to the formula *false*.)

If ξ is of the form $\neg\xi'$ then, by the induction hypothesis, ξ' is equivalent to the disjunction of the atomic descriptions in $\mathcal{A}^{\Psi}_{\xi'}$. It is easy to see that ξ is the disjunction of the atomic descriptions in $\mathcal{A}^{\Psi}_{\neg\xi'} = \mathcal{A}^{\Psi}_{M,\mathcal{X}} - \mathcal{A}^{\Psi}_{\xi'}$.

Finally, we consider the case that ξ is of the form $\exists y\, \xi'$. Recall that $M \geq d(\xi) + |\mathcal{C}| + |\mathcal{X}|$. Since $d(\xi') = d(\xi) - 1$, it is also the case that $M \geq d(\xi') + |\mathcal{C}| + |\mathcal{X} \cup \{y\}|$. By the induction hypothesis, ξ' is therefore equivalent to the disjunction of the atomic descriptions in $\mathcal{A}^{\Psi}_{\xi'}$. Clearly ξ is equivalent to $\exists y \vee_{\psi \in \mathcal{A}^{\Psi}_{\xi'}} \psi$, and standard first-order reasoning shows that $\exists y \vee_{\psi \in \mathcal{A}^{\Psi}_{\xi'}} \psi$ is equivalent to $\vee_{\psi \in \mathcal{A}^{\Psi}_{\xi'}} \exists y\, \psi$. Since $\mathcal{A}^{\Psi}_{\xi'} \subseteq \mathcal{A}^{\Psi}_{M,\mathcal{X} \cup \{y\}}$, it suffices to show that for each atomic description $\psi \in \mathcal{A}^{\Psi}_{M,\mathcal{X} \cup \{y\}}$, $\exists y\, \psi$ is equivalent to an atomic description in $\mathcal{A}^{\Psi}_{M,\mathcal{X}}$.

Consider some $\psi \in \mathcal{A}^{\Psi}_{M,\mathcal{X} \cup \{y\}}$; we can clearly pull out of the scope of $\exists y$ all the conjuncts in ψ that do not involve y. It follows that $\exists y\, \psi$ is equivalent to $\psi' \wedge \exists y\, \psi''$, where ψ'' is a conjunction of $A(y)$, where A is an atom over $\mathcal{P}$, and formulas of the form $y = z$ and $y \neq z$. It is easy to see that ψ' is a consistent atomic description over Ψ and $\mathcal{X}$ of size M. To complete the proof, we now show that $\psi' \wedge \exists y\, \psi''$ is equivalent to

ψ'. There are two cases to consider. First suppose that ψ'' contains a conjunct of the form $y = z$. Let $\psi''[y/z]$ be the result of replacing all free occurrences of y in ψ'' by z. Standard first-order reasoning (using the fact that $\psi''[y/z]$ has no free occurrences of y) shows that $\psi''[y/z]$ is equivalent to $\exists y\, \psi''[y/z]$, which is equivalent to $\exists y\, \psi''$. Since ψ is a complete atomic description which is consistent with ψ'', it follows that each conjunct of $\psi''[y/z]$ (except $z = z$) must be a conjunct of ψ', so ψ' implies $\psi''[y/z]$. It immediately follows that ψ' is equivalent to $\psi' \wedge \exists y\, \psi''$ in this case. Now suppose that there is no conjunct of the form $y = z$ in ψ''. In this case, $\exists y\, \psi''$ is certainly true if there exists a domain element satisfying atom A different from the denotations of all the symbols in $\mathcal{X} \cup \mathcal{C}$. Notice that ψ implies that there exists such an element, namely, the denotation of y. However, ψ' must already imply the existence of such an element since ψ' must force there to be enough elements satisfying A to guarantee the existence of such an element. (We remark that it is crucial for this last part of the argument that $M \geq |\mathcal{X}| + 1 + |\mathcal{C}|$.) Thus, we again have that ψ' is equivalent to $\psi' \wedge \exists y\, \psi''$. It follows that $\exists y\, \psi$ is equivalent to a consistent atomic description in $\mathcal{A}^{\Psi}_{M,\mathcal{X}}$, namely ψ', as required. ☐

For the remainder of this paper we will be interested in sentences. Thus, we restrict attention to atomic descriptions over Ψ and the empty set of variables. Moreover, we assume that all formulas mentioned are in fact sentences, and have no free variables.

Definition 3.8. *For $\Psi = \mathcal{P} \cup \mathcal{C}$, and a sentence $\xi \in \mathcal{L}(\Psi)$, we define $\mathcal{A}^{\Psi}_{\xi}$ to be the set of consistent atomic descriptions of size $d(\xi) + |\mathcal{C}|$ over Ψ such that ξ is equivalent to the disjunction of the atomic descriptions in $\mathcal{A}^{\Psi}_{\xi}$.* ☐

It will be useful for our later results to prove a simpler analogue of Theorem 3.7 for the case where the sentence ξ does not use equality or constant symbols. A *simplified atomic description over $\mathcal{P}$* is simply a size description with bound 1. Thus, it consists of a conjunction of $2^{|\mathcal{P}|}$ formulas of the form $\exists^{\geq 1} x\, A_i(x)$ or $\exists^{0} x\, A_i(x)$, one for each atom over $\mathcal{P}$. Using the same techniques as those of Theorem 3.7, we can prove the following theorem.

Theorem 3.9. *If $\xi \in \mathcal{L}^{-}(\mathcal{P})$, then ξ is equivalent to a disjunction of simplified atomic descriptions over $\mathcal{P}$.*

Proof. The proof is left to the reader. ☐

3.2. Named elements and model descriptions.

Recall that we are attempting to divide the worlds satisfying θ into classes such that:

- φ is uniform in each class, and
- the relative weight of the classes is easily computed.

In the previous section, we defined the concept of atomic description, and showed that a sentence $\theta \in \mathcal{L}(\Psi)$ is equivalent to some disjunction of atomic descriptions. This suggests that atomic descriptions might be used to classify models of θ. Līogon'kīĭ [31] has shown that this is indeed a successful approach, as long as we consider languages without constants and condition only on sentences that do not use equality. In Theorem 3.9 we showed that, for such languages, each sentence is equivalent to the disjunction of simplified atomic descriptions. The following theorem, due to Līogon'kīĭ, says that classifying models according to which simplified atomic description they satisfy leads to the desired uniformity property. This result will be a corollary of a more general theorem that we prove later.

Proposition 3.10. [31] *Suppose that $\mathcal{C} = \emptyset$. If $\varphi \in \mathcal{L}(\Phi)$ and ψ is a consistent simplified atomic description over $\mathcal{P}$, then $\Pr^{w}_{\infty}(\varphi \mid \psi)$ is either 0 or 1.*

If $\mathcal{C} \neq \emptyset$, then we do not get an analogue to Proposition 3.10 if we simply partition the worlds according to the atomic description they satisfy. For example, consider the atomic description ψ_* from Example 3.5, and the sentence $\varphi = R(c_1, c_1)$ for some binary predicate R. Clearly, by symmetry, $\Pr^w_\infty(\varphi \mid \psi_*) = 1/2$, and therefore φ is not uniform over the worlds satisfying ψ_*. We do not even need to use constant symbols, such as c_1, to construct such counterexamples. Recall that the size description in ψ_* included the conjunct $\exists^1 x\, A_1(x)$. So if $\varphi' = \exists x\, (A_1(x) \wedge R(x,x))$ then we also get $\Pr^w_\infty(\varphi' \mid \psi_*) = 1/2$.

The general problem is that, given ψ_*, φ can refer "by name" to certain domain elements and thus its truth can depend on their properties. In particular, φ can refer to domain elements that are denotations of constants in $\mathcal{C}$ as well as to domain elements that are the denotations of the "fixed-size" atoms—those atoms whose size is fixed by the atomic description. In the example above, we can view "the x such that $A_1(x)$" as a name for the unique domain element satisfying atom A_1. In any model of ψ_*, we call the denotations of the constants and elements of the fixed-size atoms the *named elements* of that model. The discussion above indicates that there is no uniformity theorem if we condition only on atomic descriptions, because an atomic expression does not fix the denotations of the nonunary predicates with respect to the named elements. This analysis suggests that we should augment an atomic description with complete information about the named elements. This leads to a finer classification of models which does have the uniformity property. To define this classification formally, we need the following definitions.

DEFINITION 3.11. *The* characteristic *of an atomic description ψ of size M is a tuple C_ψ of the form $\langle (f_1, g_1), \ldots, (f_{2^{|\mathcal{P}|}}, g_{2^{|\mathcal{P}|}}) \rangle$, where*
- *$f_i = m$ if exactly $m < M$ domain elements satisfy A_i according to ψ,*
- *$f_i = *$ if at least M domain elements satisfy A_i according to ψ,*
- *g_i is the number of distinct domain elements which are interpretations of elements in $\mathcal{C}$ that satisfy A_i according to ψ.* □

Note that we can compute the characteristic of ψ immediately from the syntactic form of ψ.

DEFINITION 3.12. *Suppose $C_\psi = \langle (f_1, g_1), \ldots, (f_{2^{|\mathcal{P}|}}, g_{2^{|\mathcal{P}|}}) \rangle$ is the characteristic of ψ. We say that an atom A_i is* active *in ψ if $f_i = *$; otherwise A_i is* passive. *Let $A(\psi)$ be the set $\{i \,:\, A_i$ is active in $\psi\}$.* □

We can now define named elements.

DEFINITION 3.13. *Given an atomic description ψ and a model $\mathcal{W}$ of ψ, the* named elements *in $\mathcal{W}$ are the elements satisfying the passive atoms and the elements that are denotations of constants.*

The number of named elements in any model of ψ is

$$\nu(\psi) = \sum_{i \in A(\psi)} g_i + \sum_{i \notin A(\psi)} f_i,$$

where $C_\psi = \langle (f_1, g_1), \ldots, (f_{2^{|\mathcal{P}|}}, g_{2^{|\mathcal{P}|}}) \rangle$, as before. □

As we have discussed, we wish to augment ψ with information about the named elements. We accomplish this using the following notion of *model fragment* which is, roughly speaking, the projection of a model onto the named elements.

DEFINITION 3.14. *Let $\psi = \sigma \wedge D$ for a size description σ and a complete description D over Ψ. A* model fragment *$\mathcal{V}$ for ψ is a model over the vocabulary Φ with domain $\{1, \ldots, \nu(\psi)\}$ such that:*
- *$\mathcal{V}$ satisfies D, and*

- *$\mathcal{V}$ satisfies the conjuncts in σ defining the sizes of the passive atoms.* □

We can now define what it means for a model $\mathcal{W}$ to satisfy a model fragment $\mathcal{V}$.

DEFINITION 3.15. *Let $\mathcal{W}$ be a model of ψ, and let $i_1, \ldots, i_{\nu(\psi)} \in \{1, \ldots, N\}$ be the named elements in $\mathcal{W}$, where $i_1 < i_2 < \cdots < i_{\nu(\psi)}$. The model $\mathcal{W}$ is said to* satisfy *the model fragment $\mathcal{V}$ if the function $F(j) = i_j$ from the domain of $\mathcal{V}$ to the domain of $\mathcal{W}$ is an isomorphism between $\mathcal{V}$ and the submodel of $\mathcal{W}$ formed by restricting to the named elements.* □

Example 3.16. Consider the atomic description ψ_* from Example 3.5. Its characteristic C_{ψ_*} is $\langle(1,0),(3,1),(*,1),(*,0)\rangle$. The active atoms are thus A_3 and A_4. Note that $g_3 = 1$ because c_2 and c_3 are constrained to denote the same element. Thus, the number of named elements $\nu(\psi_*)$ in a model of ψ_* is $1+3+1 = 5$. Therefore each model fragment for ψ_* will have domain $\{1,2,3,4,5\}$. The elements in the domain will be the named elements; these correspond to the single element in A_1, the three elements in A_2, and the unique element denoting both c_2 and c_3 in A_3.

Let Φ be $\{P_1, P_2, c_1, c_2, c_3, R\}$ where R is a binary predicate symbol. One possible model fragment $\mathcal{V}_*$ for ψ_* over Φ gives the symbols in Φ the following interpretation:

$$\begin{array}{lll} c_1^{\mathcal{V}_*} = 4, & c_2^{\mathcal{V}_*} = 3, & c_3^{\mathcal{V}_*} = 3, \\ P_1^{\mathcal{V}_*} = \{1,2,4,5\}, & P_2^{\mathcal{V}_*} = \{1,3\}, & R^{\mathcal{V}_*} = \{(1,3),(3,4)\}. \end{array}$$

It is easy to verify that $\mathcal{V}_*$ satisfies the properties of the constants as prescribed by the description D in ψ_* as well as the two conjuncts $\exists^1 x\, A_1(x)$ and $\exists^3 x\, A_2(x)$ in the size description in ψ_*.

Now, let $\mathcal{W}$ be a world satisfying ψ_*, and assume that the named elements in $\mathcal{W}$ are $3, 8, 9, 14, 17$. Then $\mathcal{W}$ satisfies $\mathcal{V}_*$ if this 5-tuple of elements has precisely the same properties in $\mathcal{W}$ as the 5-tuple $1,2,3,4,5$ does in $\mathcal{V}_*$. □

Although a model fragment is a semantic structure, the definition of satisfaction just given also allows us to regard it as a logical assertion that is true or false in any model over Φ whose domain is a subset of the natural numbers. In the following, we use this view of a model description as an assertion frequently. In particular, we freely use assertions which are the conjunction of an ordinary first-order ψ and a model fragment $\mathcal{V}$, even though the result is not a first-order formula. Under this viewpoint it makes perfect sense to use an expression such as $\Pr_\infty^w(\varphi \mid \psi \wedge \mathcal{V})$.

DEFINITION 3.17. *A* model description *augmenting ψ over the vocabulary Φ is a conjunction of ψ and a model fragment $\mathcal{V}$ for ψ over Φ. Let $\mathcal{M}^\Phi(\psi)$ be the set of model descriptions augmenting ψ. (If Φ is clear from context, we omit the subscript and write $\mathcal{M}(\psi)$ rather than $\mathcal{M}^\Phi(\psi)$.)* □

It should be clear that model descriptions are mutually exclusive and exhaustive. Moreover, as for atomic descriptions, each unary sentence θ is equivalent to some disjunction of model descriptions. From this, and elementary probability theory, we conclude the following fact, which forms the basis of our techniques for computing asymptotic conditional probabilities.

PROPOSITION 3.18. *For any $\varphi \in \mathcal{L}(\Phi)$ and $\theta \in \mathcal{L}(\Psi)$*

$$\Pr_\infty^w(\varphi \mid \theta) = \sum_{\psi \in \mathcal{A}_\theta^\Psi} \sum_{(\psi \wedge \mathcal{V}) \in \mathcal{M}(\psi)} \Pr_\infty^w(\varphi \mid \psi \wedge \mathcal{V}) \cdot \Pr_\infty^w(\psi \wedge \mathcal{V} \mid \theta),$$

if all limits exist.

As we show in the next section, model descriptions have the uniformity property so the first term in the product will always be either 0 or 1.

It might seem that the use of model fragments is a needless complication and that any model fragment, in its role as a logical assertion, will be equivalent to some first-order sentence. Consider the following definition.

DEFINITION 3.19. *Let $n = \nu(\psi)$. The* complete description capturing $\mathcal{V}$, *denoted $D_{\mathcal{V}}$, is a formula that satisfies the following:*[7]

- *$D_{\mathcal{V}}$ is a complete description over Φ and the variables $\{x_1, \ldots, x_n\}$ (see Definition 3.1),*
- *for each $i \neq j$, $D_{\mathcal{V}}$ contains a conjunct $x_i \neq x_j$, and*
- *$\mathcal{V}$ satisfies $D_{\mathcal{V}}$ when i is assigned to x_i for each $i = 1, \ldots, n$.* ◻

Example 3.20. The complete description $D_{\mathcal{V}_*}$ capturing the model fragment $\mathcal{V}_*$ from the previous example has conjuncts such as $P_1(x_1)$, $\neg P_1(x_3)$, $R(x_1, x_3)$, $\neg R(x_1, x_2)$, and $x_4 = c_1$. ◻

The distinction between a model fragment and the complete description capturing it is subtle. Clearly if a model satisfies $\mathcal{V}$, then it also satisfies $\exists x_1, \ldots, x_n\, D_{\mathcal{V}}$. The converse is not necessarily true. A model fragment places additional constraints on which domain elements are denotations of the constants and passive atoms. For example, a model fragment might entail that, in any model over the domain $\{1, \ldots, N\}$, the denotation of constant c_1 is less than that of c_2. Clearly, no first-order sentence can assert this. The main implication of this difference is combinatorial; it turns out that counting model fragments (rather than the complete descriptions that capture them) simplifies many computations considerably. Although we typically use model fragments, there are occasions where it is important to remain within first-order logic and use the corresponding complete descriptions instead. For instance, this is the case in the next subsection. Whenever we do this we will appeal to the following result, which is easy to prove.

PROPOSITION 3.21. *For any $\varphi \in \mathcal{L}(\Phi)$ and model description $\psi \wedge \mathcal{V}$ over Φ, we have*

$$\Pr_\infty^w(\varphi \mid \psi \wedge \mathcal{V}) = \Pr_\infty^w(\varphi \mid \psi \wedge \exists x_1, \ldots, x_{\nu(\psi)}\, D_{\mathcal{V}}).$$

Proof. The proof is left to the reader. ◻

3.3. A conditional 0-1 law. In the previous section, we showed how to partition θ into model descriptions. We now show that φ is uniform over each model description. That is, for any sentence $\varphi \in \mathcal{L}(\Phi)$ and any model description $\psi \wedge \mathcal{V}$, the probability $\Pr_\infty^w(\varphi \mid \psi \wedge \mathcal{V})$ is either 0 or 1. The technique we use to prove this is a generalization of Fagin's proof of the 0-1 law for first-order logic without constant or function symbols [13]. This result states that if φ is a first-order sentence in a vocabulary without constant or function symbols, then $\Pr_\infty^w(\varphi)$ is either 0 or 1.[8] It is well known that we can get asymptotic probabilities that are neither 0 nor 1 if we use constant symbols, or if we look at general conditional probabilities. However, in the special case where we condition on a model description there is still a 0-1 law. Throughout this section let $\psi \wedge \mathcal{V}$ be a fixed model description with at least one active atom, and let $n = \nu(\psi)$ be the number of named elements according to ψ.

As we said earlier, the proof of our 0-1 law is based on Fagin's proof. Like Fagin, our strategy involves constructing a theory T which, roughly speaking, states that

[7]Note that there will, in general, be more than one complete description capturing $\mathcal{V}$. We choose one of them arbitrarily for $D_{\mathcal{V}}$.

[8]As we noted in the introduction, the 0-1 law was first proved by Glebskiĭ et al. [18]. However, it is Fagin's proof technique that we are using here.

any finite fragment of a model can be extended to a larger fragment in all possible ways. We then prove two propositions.

1. T is complete; that is, for each $\varphi \in \mathcal{L}(\Phi)$, either $T \models \varphi$ or $T \models \neg\varphi$. This result, in the case of the original 0-1 law, is due to Gaifman [16].
2. For any $\varphi \in \mathcal{L}(\Phi)$, if $T \models \varphi$ then $\Pr^w_\infty(\varphi \mid \psi \wedge \mathcal{V}) = 1$.

Using the first proposition, for any sentence φ, either $T \models \varphi$ or $T \models \neg\varphi$. Therefore, using the second proposition, either $\Pr^w_\infty(\varphi \mid \psi \wedge \mathcal{V}) = 1$ or $\Pr^w_\infty(\neg\varphi \mid \psi \wedge \mathcal{V}) = 1$. The latter case immediately implies that $\Pr^w_\infty(\varphi \mid \psi \wedge \mathcal{V}) = 0$. Thus, these two propositions suffice to prove the conditional 0-1 law.

We begin by defining several concepts which will be useful in defining the theory T.

Definition 3.22. *Let $\mathcal{X}' \supseteq \mathcal{X}$, let D be a complete description over Φ and $\mathcal{X}$, and let D' be a complete description over Φ and $\mathcal{X}'$. We say that D'* extends *D if every conjunct of D is a conjunct of D'.* □

The core of the definition of T is the concept of an *extension axiom*, which asserts that any finite substructure can be extended to a larger structure containing one more element.

Definition 3.23. *Let $\mathcal{X} = \{x_1, \ldots, x_j\}$ for some k, let D be a complete description over Φ and $\mathcal{X}$, and let D' be any complete description over Φ and $\mathcal{X} \cup \{x_{j+1}\}$ that extends D. The sentence*

$$\forall x_1, x_2, \ldots, x_j\ (D \Rightarrow \exists x_{j+1} D')$$

is an extension axiom. □

In the original 0-1 law, Fagin considered the theory consisting of all the extension axioms. In our case, we must consider only those extension axioms whose components are consistent with ψ, and which extend $D_{\mathcal{V}}$.

Definition 3.24. *Given $\psi \wedge \mathcal{V}$, we define T to consist of $\psi \wedge \exists x_1, \ldots, x_n\, D_{\mathcal{V}}$ together with all extension axioms*

$$\forall x_1, x_2, \ldots, x_j\ (D \Rightarrow \exists x_{j+1} D')$$

in which D (and hence D') extends $D_{\mathcal{V}}$ and in which D' (and hence D) is consistent with ψ. □

We have used $D_{\mathcal{V}}$ rather than $\mathcal{V}$ in this definition so that T is a first-order theory. Note that the consistency condition above is not redundant, even given that the components of an extension axiom extend $D_{\mathcal{V}}$. However, inconsistency can arise only if D' asserts the existence of a new element in some passive atom (because this would contradict the size description in ψ).

We now prove the two propositions that imply the 0-1 law.

Proposition 3.25. *The theory T is complete.*

Proof. The proof is based on a result of Łoś and Vaught [40] which says that any first-order theory with no finite models, such that all of its countable models are isomorphic, is complete. The theory T obviously has no finite models. The fact that all of its countable models are isomorphic follows by a standard "back and forth" argument. That is, let $\mathcal{U}$ and $\mathcal{U}'$ be countable models of T. Without loss of generality, assume that both models have the same domain $\mathcal{D} = \{1, 2, 3, \ldots\}$. We must find a mapping $F : \mathcal{D} \rightarrow \mathcal{D}$ which is an isomorphism between $\mathcal{U}$ and $\mathcal{U}'$ with respect to Φ.

We first map the named elements in both models to each other, in the appropriate way. Recall that T contains the assertion $\exists x_1, \ldots, x_n\, D_{\mathcal{V}}$. Since $\mathcal{U} \models T$, there must exist domain elements $d_1, \ldots, d_n \in \mathcal{D}$ that satisfy $D_{\mathcal{V}}$ under the model $\mathcal{U}$. Similarly, there must exist corresponding elements $d'_1, \ldots, d'_n \in \mathcal{D}$ that satisfy $D_\mathcal{V}$ under the model $\mathcal{U}'$. We define the mapping F so that $F(d_i) = d'_i$ for $i = 1, \ldots, n$. Since $D_\mathcal{V}$ is a complete description over these elements, and the two substructures both satisfy $D_\mathcal{V}$, they are necessarily isomorphic.

In the general case, assume we have already defined F over some j elements $\{d_1, d_2, \ldots, d_j\} \in \mathcal{D}$ so that the substructure of $\mathcal{U}$ over $\{d_1, \ldots, d_j\}$ is isomorphic to the substructure of $\mathcal{U}'$ over $\{d'_1, \ldots, d'_j\}$, where $d'_i = F(d_i)$ for $i = 1, \ldots, j$. Because both substructures are isomorphic there must be a description D that is satisfied by both. Since we began by creating a mapping between the named elements, we can assume that D extends $D_\mathcal{V}$. We would like to extend the mapping F so that it eventually exhausts both domains. We accomplish this by using the even rounds of the construction (the rounds where j is even) to ensure that $\mathcal{U}$ is covered, and the odd rounds to ensure that $\mathcal{U}'$ is covered. More precisely, if j is even, let d be the first element in $\mathcal{D}$ which is not in $\{d_1, \ldots, d_j\}$. There is a model description D' extending D that is satisfied by $d_1, \ldots, d_j, d$ in $\mathcal{U}$. Consider the extension axiom in T asserting that any j elements satisfying D can be extended to $j+1$ elements satisfying D'. Since $\mathcal{U}'$ satisfies this axiom, there exists an element d' in $\mathcal{U}'$ such that $d'_1, \ldots, d'_j, d'$ satisfy D'. We define $F(d) = d'$. It is clear that the substructure of $\mathcal{U}$ over $\{d_1, \ldots, d_j, d\}$ is isomorphic to the substructure of $\mathcal{U}'$ over $\{d'_1, \ldots, d'_j, d'\}$. If j is odd, we follow the same procedure, except that we find a counterpart to the first domain element (in $\mathcal{U}'$) which does not yet have a pre-image in $\mathcal{U}$. It is easy to see that the final mapping F is an isomorphism between $\mathcal{U}$ and $\mathcal{U}'$. ☐

PROPOSITION 3.26. *For any* $\varphi \in \mathcal{L}(\Phi)$*, if* $T \models \varphi$ *then* $\Pr^w_\infty(\varphi \mid \psi \wedge \mathcal{V}) = 1$.

Proof. We begin by proving the claim for a sentence $\xi \in T$. By the construction of T, ξ is either $\psi \wedge \exists x_1, \ldots, x_n\, D_\mathcal{V}$ or an extension axiom. In the first case, Proposition 3.21 trivially implies that $\Pr^w_\infty(\xi \mid \psi \wedge \mathcal{V}) = 1$. The proof for the case that ξ is an extension axiom is based on a straightforward combinatorial argument, which we briefly sketch. Recall that one of the conjuncts of ψ is a size description σ. The sentence σ includes two types of conjuncts: those of the form $\exists^m x\, A(x)$ and those of the form $\exists^{\geq M} x\, A(x)$. Let σ' be σ with the conjuncts of the second type removed. Let ψ' be the same as ψ except that σ' replaces σ. It is easy to show that $\Pr^w_\infty(\exists^{\geq M} x\, A(x) \mid \psi' \wedge \mathcal{V}) = 1$ for any active atom A, and so $\Pr^w_\infty(\psi \mid \psi' \wedge \mathcal{V}) = 1$. Since $\psi \Rightarrow \psi'$, by straightforward probabilistic arguments, it suffices to show that $\Pr^w_\infty(\xi \mid \psi' \wedge \mathcal{V}) = 1$.

Suppose ξ is an extension axiom involving D and D', where D is a complete description over $\mathcal{X} = \{x_1, \ldots, x_j\}$ and D' is a description over $\mathcal{X} \cup \{x_{j+1}\}$ that extends D. Fix a domain size N, and some particular j domain elements $d_1, \ldots, d_j$ that satisfy D. Observe that, since D extends $D_\mathcal{V}$, all the named elements are among $d_1, \ldots, d_j$. For a given $d \notin \{d_1, \ldots, d_j\}$, let $B(d)$ denote the event that $d_1, \ldots, d_j, d$ satisfies D', conditioned on $\psi' \wedge \mathcal{V}$. The probability of $B(d)$, given that $d_1, \ldots, d_j$ satisfies D, is typically very small but is bounded away from 0 by some β independent of N. To see this, note that D' is consistent with $\psi \wedge \mathcal{V}$ (because D' is part of an extension axiom) and so there is a consistent way of choosing how d is related to $d_1, \ldots, d_j$ so as to satisfy D'. Then observe that the total number of possible ways to choose d's properties (as they relate to $d_1, \ldots, d_j$) is independent of N. Since D extends $D_\mathcal{V}$, the model fragment defined over the elements $d_1, \ldots, d_j$ satisfies $\psi' \wedge \mathcal{V}$. (Note that it does not necessarily satisfy ψ, which is why we replaced ψ with ψ'.) Since the properties of an element d and its relation to $d_1, \ldots, d_j$ can be chosen independently of the properties of a different element

d', the different events $B(d), B(d'), \ldots$ are all independent. Therefore, the probability that there is no domain element at all that, together with $d_1, \ldots, d_j$, satisfies D' is at most $(1-\beta)^{N-j}$. This bounds the probability of the extension axiom being false, relative to fixed $d_1, \ldots, d_j$. There are exactly $\binom{N}{j}$ ways of choosing j elements, so the probability of the axiom being false anywhere in a model is at most $\binom{N}{j}(1-\beta)^{N-j}$. However, this tends to 0 as N goes to infinity. Therefore, the axiom $\forall x_1, \ldots, x_j\,(D \Rightarrow \exists x_{j+1}\, D')$ has asymptotic probability 1 given $\psi' \wedge \mathcal{V}$, and therefore also given $\psi \wedge \mathcal{V}$.

It remains to deal only with the case of a general formula $\varphi \in \mathcal{L}(\Phi)$ such that $T \models \varphi$. By the compactness theorem for first-order logic, if $T \models \varphi$ then there is some finite conjunction of assertions $\xi_1, \ldots, \xi_m \in T$ such that $\wedge_{i=1}^m \xi_i \models \varphi$. By the previous case, each such ξ_i has asymptotic probability 1, and therefore so does this finite conjunction. Hence, the asymptotic probability $\Pr_\infty^w(\varphi \mid \psi \wedge \mathcal{V})$ is also 1. ∎

As outlined above, this concludes the proof of the main theorem of this section, which we now state.

THEOREM 3.27. *For any sentence $\varphi \in \mathcal{L}(\Phi)$ and model description $\psi \wedge \mathcal{V}$, $\Pr_\infty^w(\varphi \mid \psi \wedge \mathcal{V})$ is either 0 or 1.*

Note that if ψ is a simplified atomic description, then there are no named elements in any model of ψ. Therefore, the only model description augmenting ψ is simply ψ itself. Thus Proposition 3.10, which is Līogon'kĭĭ's result, is a corollary of the above theorem.

3.4. Computing the relative weights of model descriptions. We now want to compute the relative weights of model descriptions. It will turn out that certain model descriptions are dominated by others, so that their relative weight is 0, while all the dominating model descriptions have equal weight. Thus, the problem of computing the relative weights of model descriptions reduces to identifying the dominating model descriptions. There are two factors that determine which model descriptions dominate. The first, and more significant, is the number of active atoms; the second is the number of named elements. Let $\alpha(\psi)$ denote the number of active atoms according to ψ.

To compute these relative weights of the model descriptions, we must evaluate $\#world_N^\Phi(\psi \wedge \mathcal{V})$. The following lemma gives a precise expression for the asymptotic behavior of this function as N grows large.

LEMMA 3.28. *Let ψ be a consistent atomic description of size $M \geq |\mathcal{C}|$ over Ψ, and let $(\psi \wedge \mathcal{V}) \in \mathcal{M}^\Phi(\psi)$.*

(a) *If $\alpha(\psi) = 0$ and $N > \nu(\psi)$, then $\#world_N^\Psi(\psi) = 0$. In particular, this holds for all $N > 2^{|\mathcal{P}|}M$.*

(b) *If $\alpha(\psi) > 0$, then*

$$\#world_N^\Phi(\psi \wedge \mathcal{V}) \sim \binom{N}{n} a^{N-n} 2^{\sum_{i \geq 2} b_i (N^i - n^i)},$$

where $a = \alpha(\psi)$, $n = \nu(\psi)$, and b_i is the number of predicates of arity i in Φ.

Proof. Suppose that $C_\psi = \langle (f_1, g_1), \ldots, (f_{2^{|\mathcal{P}|}}, g_{2^{|\mathcal{P}|}}) \rangle$ is the characteristic of ψ. Let $\mathcal{W}$ be a model of cardinality N, and let N_i be the number of domain elements in $\mathcal{W}$ satisfying atom A_i. In this case, we say that the *profile of* $\mathcal{W}$ is $\langle N_1, \ldots, N_{2^{|\mathcal{P}|}} \rangle$. Clearly we must have $N_1 + \cdots + N_{2^{|\mathcal{P}|}} = N$. We say that the profile $\langle N_1, \ldots, N_{2^{|\mathcal{P}|}} \rangle$ is *consistent* with C_ψ if $f_i \neq *$ implies that $N_i = f_i$, while $f_i = *$ implies that $N_i \geq M$. Notice that if $\mathcal{W}$ is a model of ψ, then the profile of $\mathcal{W}$ must be consistent with C_ψ.

For part (a), observe that if $\alpha(\psi) = 0$ and $N > \sum_i f_i$, then there can be no models of cardinality N whose profile is consistent with C_ψ. However, if $\alpha(\psi) = 0$, then $\sum_i f_i$ is precisely $\nu(\psi)$. Hence there can be no models of ψ of cardinality N if $N > \nu(\psi)$. Moreover, since $\nu(\psi) \leq 2^{|\mathcal{P}|}M$, the result holds for any $N > 2^{|\mathcal{P}|}$. This proves part (a).

For part (b), let us first consider how many ways there are of choosing a world satisfying $\psi \wedge \mathcal{V}$ with cardinality N and profile $\langle N_1, \ldots, N_{2^{|\mathcal{P}|}}\rangle$. To do the count, we first choose which elements are to be the named elements in the domain. Clearly, there are $\binom{N}{n}$ ways in which this can be done. Once we choose the named elements, their properties are completely determined by $\mathcal{V}$.

It remains to specify the rest of the properties of the world. Let R be a nonunary predicate of arity $i \geq 2$. To completely describe the behavior of R in a world, we need to specify which of the N^i i-tuples over the domain are in the denotation of R. We have already specified this for those i-tuples all of whose components are named elements. There are n^i such i-tuples. Therefore, we have $N^i - n^i$ i-tuples left to specify. Since each subset is a possible denotation, we have $2^{N^i - n^i}$ possibilities for the denotation of R. The overall number of choices for the denotations of all nonunary predicates in the vocabulary is therefore $2^{\Sigma_{i\geq 2} b_i(N^i - n^i)}$.

It remains only to choose the denotations of the unary predicates for the $N' = N - n$ domain elements that are not named. Let $i_1, \ldots, i_a$ be the active atoms in ψ, and let $h_j = N_{i_j} - g_{i_j}$ for $j = 1, \ldots, a$. Thus, we need to compute all the ways of partitioning the remaining N' elements so that there are h_j elements satisfying atom A_{i_j}; there are $\binom{N'}{h_1\ h_2\ \ldots\ h_a}$ ways of doing this.

We now need to sum over all possible profiles, i.e., those consistent with $\psi \wedge \mathcal{V}$. If $i_j \in A(\psi)$, then there must be at least M domain elements satisfying A_{i_j}. Therefore $N_{i_j} \geq M$, and $h_j = N_{i_j} - g_{i_j} \geq M - g_{i_j}$. This is the only constraint on h_j. Thus, it follows that

$$\#world^{\Phi}_N(\psi \wedge \mathcal{V}) \sim \sum_{\{h_1,\ldots,h_a:\ h_1+\cdots+h_a = N',\ \forall j\ h_j \geq M - g_{i_j}\}} \binom{N}{n} 2^{\sum_{i\geq 2} b_i(N^i - n^i)} \binom{N'}{h_1\ \ldots\ h_a}.$$

This is equal to

$$\binom{N}{n} 2^{\sum_{i\geq 2} b_i(N^i - n^i)} S$$

for

$$S = \sum_{\{h_1,\ldots,h_a:\ h_1+\cdots+h_a = N',\ \forall j\ h_j \geq M - g_{i_j}\}} \binom{N'}{h_1\ \ldots\ h_a}.$$

It remains to get a good asymptotic estimate for S. Notice that

$$\sum_{\{h_1,\ldots,h_a:\ h_1+\cdots+h_a = N'\}} \binom{N'}{h_1\ \ldots\ h_a} = a^{N'},$$

since the sum can be viewed as describing all possible ways to assign one of a possible atoms to each of N' elements. Our goal is to show that $a^{N'}$ is actually a good approximation for S as well. Clearly $S < a^{N'}$. Let

$$S_j = \sum_{\{h_1,\ldots,h_a:\ h_j < M,\ h_1+\cdots+h_a = N'\}} \binom{N'}{h_1\ \ldots\ h_a}.$$

Straightforward computation shows that

$$\begin{aligned}
S_1 &= \sum_{\{h_1,\ldots,h_a:\ h_1<M,\ h_1+\cdots+h_a=N'\}} \binom{N'}{h_1\ \ldots\ h_a} \\
&= \sum_{h_1=0}^{M-1} \sum_{\{h_2,\ldots,h_a:\ h_2+\cdots+h_a=N'-h_1\}} \binom{N'}{h_1}\binom{N'-h_1}{h_2\ \ldots\ h_a} \\
&\le \sum_{h_1=0}^{M-1} \frac{(N')^{h_1}}{h_1!}(a-1)^{N'-h_1} \\
&< MN^M(a-1)^{N'}\ .
\end{aligned}$$

Similar arguments show that $S_j < MN^M(a-1)^{N'}$ for all j. It follows that

$$\begin{aligned}
S &> \sum_{\{h_1,\ldots,h_a:h_1+\cdots+h_a=N'\}} \binom{N'}{h_1\ \ldots\ h_a} - (S_1+\cdots+S_a) \\
&> a^{N'} - aMN^M(a-1)^{N'}\ .
\end{aligned}$$

Therefore,

$$S \sim a^{N'} = a^{N-n},$$

thus concluding the proof. □

The asymptotic behavior described in this lemma motivates the following definition.

DEFINITION 3.29. *Given an atomic description ψ over Ψ, let the* degree *of ψ, written $\Delta(\psi)$, be the pair $(\alpha(\psi),\nu(\psi))$, and let degrees be ordered lexicographically. We extend this definition to sentences as follows. For $\theta \in \mathcal{L}(\Psi)$, we define the* degree of θ over Ψ, *written $\Delta^\Psi(\theta)$, to be* $\max_{\psi\in\mathcal{A}_\theta^\Psi}\Delta(\psi)$, *and the* activity count *of θ to be $\alpha^\Psi(\theta)$ (i.e., the first component of $\Delta^\Psi(\theta)$).* □

One important conclusion of this lemma justifies our treatment of well-definedness (Definition 2.2) when conditioning on unary formulas. It shows that if θ is satisfied in some "sufficiently large" model, then it is satisfiable over all "sufficiently large" domains.

LEMMA 3.30. *Suppose that $\theta \in \mathcal{L}(\Psi)$, and $M = d(\theta)+|\mathcal{C}_\theta|$. Then the following conditions are equivalent:*

(a) *θ is satisfied in some model of cardinality greater than $2^{|\mathcal{P}|}M$,*

(b) $\alpha^\Psi(\theta) > 0$,

(c) *for all $N > 2^{|\mathcal{P}|}M$, θ is satisfiable in some model of cardinality N,*

(d) *$\Pr_\infty^w(* \mid \theta)$ is well defined.*

Proof. By definition, θ is satisfiable in some model of cardinality N iff $\#world_N^\Psi(\theta) > 0$. We first show that (a) implies (b). If θ is satisfied in some model of cardinality N greater than $2^{|\mathcal{P}|}M$, then there is some atomic description $\psi\in\mathcal{A}_\theta^\Psi$ such that ψ is satisfied in some model of cardinality N. Using part (a) of Lemma 3.28, we deduce that $\alpha(\psi) > 0$ and therefore that $\alpha^\Psi(\theta) > 0$. That (b) entails (c) can be verified by examining the proof of Lemma 3.28. That (c) implies (d) and (d) implies (a) is immediate from the definition of well-definedness. □

For the case of sentences in the languages without equality or constants, the condition for well-definedness simplifies considerably.

COROLLARY 3.31. *If $\theta \in \mathcal{L}^-(\mathcal{P})$, then $\Pr^w_\infty(* \mid \theta)$ is well defined iff θ is satisfiable.*

Proof. The only if direction is obvious. For the other, if θ is consistent, then it is equivalent to a nonempty disjunction of consistent simplified atomic descriptions. Any consistent simplified atomic description has arbitrarily large models. ∎

We remark that we can extend our proof techniques to show that Corollary 3.31 holds even if $\mathcal{C} \neq \emptyset$, although we must still require that θ does not mention equality. We omit details here.

For the remainder of this paper, we will consider only sentences θ such that $\alpha^\Psi(\theta) > 0$.

Lemma 3.28 shows that, asymptotically, the number of worlds satisfying $\psi \wedge \mathcal{V}$ is completely determined by the degree of ψ. Model descriptions of higher degree have many more worlds, and therefore dominate. On the other hand, model descriptions with the same degree have the same number of worlds at the limit, and are therefore equally likely. This observation allows us to compute the relative weights of different model descriptions.

DEFINITION 3.32. *For any degree $\delta = (a, n)$, let $\mathcal{A}^{\Psi,\delta}_\theta$ be the set of atomic descriptions $\psi \in \mathcal{A}^\Psi_\theta$ such that $\Delta(\psi) = \delta$. For any set of atomic descriptions $\mathcal{A}'$, we use $\mathcal{M}(\mathcal{A}')$ to denote $\cup_{\psi \in \mathcal{A}'} \mathcal{M}(\psi)$.* ∎

THEOREM 3.33. *Let $\theta \in \mathcal{L}(\Psi)$ and $\Delta^\Psi(\theta) = \delta \geq (1, 0)$. Let ψ be an atomic description in $\mathcal{A}^\Psi_\theta$, and let $\psi \wedge \mathcal{V} \in \mathcal{M}^\Phi(\psi)$.*

(a) *If $\Delta(\psi) < \delta$ then $\Pr^w_\infty(\psi \wedge \mathcal{V} \mid \theta) = 0$.*

(b) *If $\Delta(\psi) = \delta$ then $\Pr^w_\infty(\psi \wedge \mathcal{V} \mid \theta) = 1/|\mathcal{M}^\Phi(\mathcal{A}^{\Psi,\delta}_\theta)|$.*

Proof. We begin with part (a). Since $\Delta^\Psi(\theta) = \delta = (a, n)$, there must exist some atomic description $\psi' \in \mathcal{A}^\Psi_\theta$ with $\Delta(\psi') = \delta$. Let $\psi' \wedge \mathcal{V}'$ be some model description in $\mathcal{M}(\psi')$.

$$\begin{aligned}
\Pr^w_N(\psi \wedge \mathcal{V} \mid \theta) &= \frac{\#world^\Phi_N(\psi \wedge \mathcal{V})}{\#world^\Phi_N(\theta)} \\
&\leq \frac{\#world^\Phi_N(\psi \wedge \mathcal{V})}{\#world^\Phi_N(\psi' \wedge \mathcal{V}')} \\
&\sim \frac{\binom{N}{\nu(\psi)} (\alpha(\psi))^{N-\nu(\psi)} 2^{\sum_{i \geq 2} b_i (N^i - \nu(\psi)^i)}}{\binom{N}{n} a^{N-n} 2^{\sum_{i\geq 2} b_i (N^i - n^i)}} \\
&= O(N^{\nu(\psi)-n} (\alpha(\psi)/a)^N).
\end{aligned}$$

The last step uses the fact that n and $\nu(\psi)$ can be considered to be constant, and that for any constant k, $\binom{N}{k} \sim N^k/k!$. Since $\Delta(\psi) < \delta = (a, n)$, either $\alpha(\psi) < a$ or $\alpha(\psi) = a$ and $\nu(\psi) < n$. In either case, it is easy to see that $N^{\nu(\psi)-n}(\alpha(\psi)/a)^N$ tends to 0 as $N \to \infty$, giving us our result.

To prove part (b), we first observe that, due to part (a), we can essentially ignore all model descriptions of low degree. That is:

$$\#world^\Phi_N(\theta) \sim \sum_{(\psi' \wedge \mathcal{V}') \in \mathcal{M}(\mathcal{A}^{\Psi,\delta}_\theta)} \#world^\Phi_N(\psi' \wedge \mathcal{V}').$$

Therefore,

$$\Pr^w_N(\psi \wedge \mathcal{V} \mid \theta) = \frac{\#world^\Phi_N(\psi \wedge \mathcal{V})}{\sum_{(\psi' \wedge \mathcal{V}') \in \mathcal{M}(\mathcal{A}^{\Psi,\delta}_\theta)} \#world^\Phi_N(\psi' \wedge \mathcal{V}')}$$

$$\sim \frac{\binom{N}{n} a^{N-n} 2^{\sum_{i\geq 2} b_i(N^i-n^i)}}{\sum_{(\psi'\wedge\mathcal{V}')\in\mathcal{M}(\mathcal{A}_\theta^{\Psi,\delta})} \binom{N}{n} a^{N-n} 2^{\sum_{i\geq 2} b_i(N^i-n^i)}}$$
$$= \frac{1}{|\mathcal{M}(\mathcal{A}_\theta^{\Psi,\delta})|},$$

as desired. □

Combining this result with Proposition 3.18, we deduce the following.

THEOREM 3.34. *For any $\varphi \in \mathcal{L}(\Phi)$ and $\theta \in \mathcal{L}(\Psi)$ such that $\Delta^\Psi(\theta) = \delta \geq (1,0)$,*

$$\Pr_\infty^w(\varphi \mid \theta) = \sum_{(\psi\wedge\mathcal{V})\in\mathcal{M}(\mathcal{A}_\theta^{\Psi,\delta})} \Pr_\infty^w(\varphi \mid \psi \wedge \mathcal{V})/|\mathcal{M}(\mathcal{A}_\theta^{\Psi,\delta})|.$$

This result, together with the techniques of the next section, will allow us to compute asymptotic conditional probabilities.

The results of Līogon'kīĭ are a simple corollary of the above theorem. For an activity count a, let $\mathcal{A}_\theta^{\Psi,a}$ denote the set of atomic descriptions $\psi \in \mathcal{A}_\theta^\Psi$ such that $\alpha(\psi) = a$.

THEOREM 3.35. [31] *Assume that $\mathcal{C} = \emptyset$, $\varphi \in \mathcal{L}(\Phi)$, $\theta \in \mathcal{L}^-(\mathcal{P})$, and $\alpha^\mathcal{P}(\theta) = a > 0$. Then $\Pr_\infty^w(\varphi \mid \theta) = \sum_{\psi\in\mathcal{A}_\theta^{\mathcal{P},a}} \Pr_\infty^w(\varphi \mid \psi)/|\mathcal{A}_\theta^{\mathcal{P},a}|$.*

Proof. By Theorem 3.9, a sentence $\theta \in \mathcal{L}^-(\mathcal{P})$ is the disjunction of the simplified atomic descriptions in $\mathcal{A}_\theta^\mathcal{P}$. A simplified atomic description ψ has no named elements, and therefore $\Delta(\psi) = (\alpha(\psi), 0)$. Moreover, $\mathcal{M}(\psi) = \{\psi\}$ for any $\psi \in \mathcal{A}_\theta^\mathcal{P}$. The result now follows trivially from the previous theorem. □

This calculation simplifies somewhat if φ and θ are both monadic. In this case, we assume without loss of generality that $d(\varphi) = d(\theta)$. (If not, we can replace φ with $\varphi \wedge \theta$ and θ with $\theta \wedge (\varphi \vee \neg\varphi)$.) This allows us to assume that $\mathcal{A}_{\varphi\wedge\theta}^\Psi \subseteq \mathcal{A}_\theta^\Psi$, thus simplifying the presentation.

COROLLARY 3.36. *Assume that $\varphi, \theta \in \mathcal{L}^-(\mathcal{P})$, and $\alpha^\mathcal{P}(\theta) = a > 0$. Then*

$$\Pr_\infty^w(\varphi \mid \theta) = \frac{|\mathcal{A}_{\varphi\wedge\theta}^{\mathcal{P},a}|}{|\mathcal{A}_\theta^{\mathcal{P},a}|}.$$

Proof. Since φ is monadic, $\varphi \wedge \theta$ is equivalent to a disjunction of the atomic descriptions $\mathcal{A}_{\varphi\wedge\theta}^\mathcal{P} \subseteq \mathcal{A}_\theta^\mathcal{P}$. Atomic descriptions are mutually exclusive; thus, for $\psi \in \mathcal{A}_\theta^\Psi$, $\Pr_\infty^w(\varphi \mid \psi) = 1$ if $\psi \in \mathcal{A}_{\varphi\wedge\theta}^\Psi$ and $\Pr_\infty^w(\varphi \mid \psi) = 0$ otherwise. The result then follows immediately from Theorem 3.35. □

3.5. Asymptotic probabilities for random structures. We now turn our attention to computing asymptotic conditional probabilities using the random-structures method. There are two cases. In the first, there is at least one nonunary predicate in the vocabulary. In this case, random structures is equivalent to random worlds, so that the results in the previous section apply without change.

THEOREM 3.37. *If $\Phi \neq \Psi$ then for any $\varphi \in \mathcal{L}(\Phi)$ and $\theta \in \mathcal{L}(\Psi)$, $\Pr_\infty^{s,\Phi}(\varphi \mid \theta) = \Pr_\infty^w(\varphi \mid \theta)$.*

Proof. Since $\Phi \neq \Psi$, there is at least one nonunary predicate in Φ that does not appear in θ. We can therefore apply Corollary 2.10, and conclude the desired result. □

Random worlds and random structures differ in the second case, when all predicates are unary, but the absence of high-arity predicates makes this a much simpler problem. For the rest of this section, we investigate the asymptotic probability of φ given θ using random structures, for $\varphi, \theta \in \mathcal{L}(\Psi)$. As discussed earlier, we can assume without loss of generality that $\mathcal{A}^{\Psi}_{\varphi\wedge\theta} \subseteq \mathcal{A}^{\Psi}_{\theta}$.

We will use the same basic technique of dividing the structures satisfying θ into classes, and computing the probability of φ on each part. In the case of random structures, however, we partition structures according to the atomic description they satisfy. That is, our computation makes use of the equation

$$\Pr^{s,\Psi}_{\infty}(\varphi \mid \theta) = \sum_{\psi \in \mathcal{A}^{\Psi}_{\theta}} \Pr^{s,\Psi}_{\infty}(\varphi \mid \psi) \Pr^{s,\Psi}_{\infty}(\psi \mid \theta).$$

As for the case of random worlds, we assign weights to atomic descriptions by counting structures. The following lemma computes $\#struct^{\Psi}_{N}(\psi)$ for an atomic description ψ. In the case of random worlds, we saw in Lemma 3.28 that certain model descriptions $\psi \wedge \mathcal{V}$ dominate others, based on the activity count $\alpha(\psi)$ and the number of named elements $\nu(\psi)$ of the atomic description. The following analogue of Lemma 3.28 shows that, for the random-structures method, atomic descriptions of higher activity count $\alpha(\psi)$ dominate regardless of the number of named elements.

LEMMA 3.38. *Let ψ be a consistent atomic description of size $M \geq |\mathcal{C}|$ over Ψ.*

(a) *If $\alpha(\psi) = 0$ and $N > \nu(\psi)$, then $\#struct^{\Psi}_{N}(\psi) = 0$. In particular, this holds for $N > 2^{|\mathcal{P}|}M$.*

(b) *If $\alpha(\psi) > 0$ then $\#struct^{\Psi}_{N}(\psi) \sim \frac{N^{\alpha(\psi)-1}}{(\alpha(\psi)-1)!}$.*

Proof. Part (a) follows immediately from Lemma 3.28(a), since $\#struct^{\Psi}_{N}(\psi) = 0$ iff $\#world^{\Psi}_{N}(\psi) = 0$.

We now proceed to show part (b). Suppose that $C_\psi = \langle (f_1, g_1), \ldots, (f_{2^{|\mathcal{P}|}}, g_{2^{|\mathcal{P}|}}) \rangle$ is the characteristic of ψ. Let $\mathcal{S}$ be a structure of cardinality N. For any of the models in $\mathcal{S}$, let N_i be the number of domain elements satisfying atom A_i (because $\mathcal{S}$ is an isomorphism class, N_i must be the same for all worlds in the class). As before, we say that the *profile of* $\mathcal{S}$ is $\langle N_1, \ldots, N_{2^{|\mathcal{P}|}} \rangle$. Clearly we must have $N_1 + \cdots + N_{2^{|\mathcal{P}|}} = N$. Recall that the profile $\langle N_1, \ldots, N_{2^{|\mathcal{P}|}} \rangle$ is consistent with C_ψ if $f_i \neq *$ implies that $N_i = f_i$, while $f_i = *$ implies that $N_i \geq M$. Notice that if $\mathcal{S}$ is a structure of ψ, then the profile of $\mathcal{S}$ must be consistent with C_ψ. In fact, there is a unique structure consistent with ψ with cardinality N and profile $\langle N_1, \ldots, N_{2^{|\mathcal{P}|}} \rangle$. This is because a structure is determined by the number of elements in each atom, the assignment of constants to atoms, and the equality relations between the constants. The first part is determined by the profile, while the second and third are determined by ψ. It therefore remains to count only the number of profiles consistent with C_ψ. Let $N' = N - \sum_{i \notin A(\psi)} f_i$, and let $i_1, \ldots, i_a$, $a = \alpha(\psi)$, be the active components of C_ψ. We want to compute

$$S = |\{\langle N_{i_1}, \ldots, N_{i_a} \rangle : \ N_{i_1} + \cdots + N_{i_a} = N', \ \forall j \ N_{i_j} \geq M\}| \ .$$

Notice that, since $\sum_{i \notin A(\psi)} f_i$ and a are constants,

$$|\{\langle N_{i_1}, \ldots, N_{i_a} \rangle : \ N_{i_1} + \cdots + N_{i_a} = N'\}| = \binom{N' + a - 1}{a - 1} \sim \frac{(N')^{a-1}}{(a-1)!} \sim \frac{N^{a-1}}{(a-1)!} \ .$$

As in the proof of Lemma 3.28, let $S_j = |\{\langle N_{i_1}, \ldots, N_{i_a} \rangle : \ N_{i_1} + \cdots + N_{i_a} = N', \ N_{i_j} < M\}|$. It is easy to see that

$$\begin{aligned} S_1 &= \sum_{N_{i_1}=0}^{M-1} |\{\langle N_{i_2}, \ldots, N_{i_a} \rangle : \ N_{i_2} + \cdots + N_{i_a} = N' - N_{i_1}\}| \\ &< M(N')^{a-2} \ , \end{aligned}$$

and similarly for all other S_j. Therefore,

$$\begin{aligned} S &\geq \binom{N'+a-1}{a-1} - (S_1 + \cdots + S_a) \\ &> \binom{N'+a-1}{a-1} - aM(N')^{a-2} \\ &\sim \frac{(N')^{a-1}}{(a-1)!} - aM(N')^{a-2} \\ &\sim \frac{N^{a-1}}{(a-1)!} \ . \end{aligned}$$

It follows that $S \sim \frac{N^{a-1}}{(a-1)!}$, as desired. ☐

COROLLARY 3.39. *If $\theta \in \mathcal{L}(\Psi)$, $\alpha^\Psi(\theta) = a > 0$, and $\psi \in \mathcal{A}_\theta^\Psi$, then*

(a) *if $\alpha(\psi) < a$ then $\Pr_\infty^{s,\Psi}(\psi \mid \theta) = 0$,*

(b) *if $\alpha(\psi) = a$ then $\Pr_\infty^{s,\Psi}(\psi \mid \theta) = 1/|\mathcal{A}_\theta^{\Psi,a}|$.*

Proof. Using Lemma 3.38, we can deduce that

$$\Pr_N^{s,\Psi}(\psi \mid \theta) \sim \frac{N^{\alpha(\psi)-1}/(\alpha(\psi)-1)!}{\sum_{\psi' \in \mathcal{A}_\theta^\Psi} N^{\alpha(\psi')-1}/(\alpha(\psi')-1)!} \ .$$

As in the proof of Theorem 3.33, we can deduce that if $\alpha(\psi) < a = \alpha^\Psi(\theta)$, then $\Pr_\infty^{s,\Psi}(\psi \mid \theta) = 0$. Therefore

$$\#struct_N^\Psi(\theta) \sim \sum_{\psi' \in \mathcal{A}_\theta^{\Psi,a}} \#struct_N^\Psi(\psi').$$

Since $\#struct_N^\Phi(\psi')$ is asymptotically the same for all ψ' with the same activity count $\alpha(\psi')$, we deduce that if $\alpha(\psi) = a$, then $\Pr_\infty^{s,\Psi}(\psi \mid \theta) = 1/|\mathcal{A}_\theta^{\Psi,a}|$. ☐

We can now complete the computation of the value of $\Pr_\infty^{s,\Psi}(\varphi \mid \theta)$ for the case of unary φ, θ.

THEOREM 3.40. *If $\varphi, \theta \in \mathcal{L}(\Psi)$ and $a = \alpha^\Psi(\theta) > 0$, then*

$$\Pr_\infty^{s,\Psi}(\varphi \mid \theta) = \frac{|\mathcal{A}_{\varphi\wedge\theta}^{\Psi,a}|}{|\mathcal{A}_\theta^{\Psi,a}|} \ .$$

Proof. Recall that

$$\Pr_\infty^{s,\Psi}(\varphi \mid \theta) = \sum_{\psi \in \mathcal{A}_\theta^\Psi} \Pr_\infty^{s,\Psi}(\varphi \mid \psi) \Pr_\infty^{s,\Psi}(\psi \mid \theta).$$

We have already computed $\Pr_\infty^{s,\Psi}(\psi \mid \theta)$. It remains to compute $\Pr_\infty^{s,\Psi}(\varphi \mid \psi)$ for an atomic description ψ. Recall that $\varphi \wedge \theta$ is equivalent to a disjunction of the atomic descriptions $\mathcal{A}_{\varphi\wedge\theta}^\Psi \subseteq \mathcal{A}_\theta^\Psi$, and that atomic descriptions are mutually exclusive. Therefore, for $\psi \in \mathcal{A}_\theta^\Psi$, it is easy to see that $\Pr_\infty^{s,\Psi}(\varphi \mid \psi) = 1$ if $\psi \in \mathcal{A}_{\varphi\wedge\theta}^\Psi$ and $\Pr_\infty^{s,\Psi}(\varphi \mid \psi) = 0$ otherwise. Since $\Pr_\infty^{s,\Psi}(\psi \mid \theta)$ is 0 except if $\psi \in \mathcal{A}_\theta^{\Psi,a}$, it follows from Corollary 3.39 that

$$\Pr_\infty^{s,\Psi}(\varphi \mid \theta) = \frac{|\mathcal{A}_{\varphi\wedge\theta}^{\Psi,a}|}{|\mathcal{A}_\theta^{\Psi,a}|},$$

as desired. □

Recall that if $\xi \in \mathcal{L}^-(\mathcal{P})$, then $\Delta^{\mathcal{P}}(\xi) = (\alpha^{\mathcal{P}}(\xi), 0)$. Thus, comparing Corollary 3.36 with Theorem 3.40 shows that, for formulas in $\mathcal{L}^-(\mathcal{P})$, random worlds and random structures are the same.

Corollary 3.41. *If $\varphi, \theta \in \mathcal{L}^-(\mathcal{P})$, then for any $\Psi \supseteq \mathcal{P}$, we have* $\Pr^w_\infty(\varphi \mid \theta) = \Pr^{s,\Psi}_\infty(\varphi \mid \theta)$.

Note that, although in general the asymptotic conditional probability in the case of random structures may depend on the vocabulary, for formulas without constant symbols or equality, it does not.

Corollary 3.42. *If $\varphi, \theta \in \mathcal{L}^-(\mathcal{P})$ and $\mathcal{P}_{\varphi\wedge\theta} \subseteq \Psi \cap \Psi'$, then* $\Pr^{s,\Psi}_\infty(\varphi \mid \theta) = \Pr^{s,\Psi'}_\infty(\varphi \mid \theta) = \Pr^w_\infty(\varphi \mid \theta)$.

4. Complexity analysis. In this section we investigate the computational complexity of problems associated with asymptotic conditional probabilities. In fact, we consider three problems: deciding whether the asymptotic probability is well defined, computing it, and approximating it. As we did in the previous section, we begin with the case of random worlds. As we shall see, the same complexity results also hold for the random-structures case (even though, as we have seen, the actual values being computed can differ between random structures and random worlds). The analysis for the unary case of random structures is given in §4.6.

Our computational approach is based on Theorem 3.34, which tells us that

$$\Pr^w_\infty(\varphi \mid \theta) = \frac{1}{|\mathcal{M}(\mathcal{A}^{\Psi,\delta}_\theta)|} \cdot \sum_{(\psi\wedge\mathcal{V})\in\mathcal{M}(\mathcal{A}^{\Psi,\delta}_\theta)} \Pr^w_\infty(\varphi \mid \psi \wedge \mathcal{V}).$$

The basic structure of the algorithms we give for computing $\Pr^w_\infty(\varphi \mid \theta)$ is simply to enumerate model descriptions $\psi \wedge \mathcal{V}$ and, for those of the maximum degree, compute the conditional probability $\Pr^w_\infty(\varphi \mid \psi \wedge \mathcal{V})$. In §4.1 we show how to compute this latter probability.

The complexity of computing asymptotic probabilities depends on several factors: whether the vocabulary is finite, whether there is a bound on the depth of quantifier nesting, whether equality is used in θ, whether nonunary predicates are used, and whether there is a bound on predicate arities. If we consider a fixed and finite vocabulary there are just two cases: if there is no bound on the depth of quantifier nesting then computing probabilities is PSPACE-complete; otherwise the computation can be done in linear time. The case in which the vocabulary is not fixed, which is the case more typically considered in complexity theory, is more complicated. The problem of computing probabilities is complete for the class #EXP (defined below) if either (a) equality is not used in θ and there is some fixed bound on the arity of predicates that can appear in φ, or (b) all predicates in φ are unary. Weakening these conditions in any way—allowing equality while maintaining any arity bound greater than one, or allowing unbounded arity even without using equality in θ—gives the same complexity as the general case (which is complete for a class we call #TA(EXP, LIN), defined later). All these results for the case of an unbounded vocabulary use formulas with quantifier depth 2. As suggested in the introduction, the complexity of the problem drops in the case of formulas of depth 1. A detailed analysis for this case can be found in [27].

4.1. Computing the 0-1 probabilities. The method we give for computing $\Pr^w_\infty(\varphi \mid \psi\wedge\mathcal{V})$ is an extension of Grandjean's algorithm [20] for computing asymptotic probabilities in the unconditional case. For the purposes of this section, fix a model description $\psi \wedge \mathcal{V}$ over Φ. In our proof of the conditional 0-1 law (§3.3), we defined a theory T corresponding to $\psi \wedge \mathcal{V}$. We showed that T is a complete and consistent theory, and that $\varphi \in \mathcal{L}(\Phi)$ has asymptotic probability 1 iff $T \models \varphi$. We therefore need an algorithm that decides whether $T \models \varphi$.

Grandjean's original algorithm decides whether $\Pr_\infty^w(\varphi)$ is 0 or 1 for a sentence φ with no constant symbols. For this case, the theory T consists of all possible extension axioms, rather than just the ones involving model descriptions extending $D_\mathcal{V}$ and consistent with ψ (see Definition 3.24). The algorithm has a recursive structure, which at each stage attempts to decide something more general than whether $T \models \varphi$. It decides whether $T \models D \Rightarrow \xi$, where

- D is a complete description over Φ and the set $\mathcal{X}_j = \{x_1, \ldots, x_j\}$ of variables, and
- $\xi \in \mathcal{L}(\Phi)$ is a formula whose only free variables (if any) are in $\mathcal{X}_j$.

The algorithm begins with $j = 0$. In this case, D is a complete description over $\mathcal{X}_0$ and Φ. Since Φ contains no constants and $\mathcal{X}_0$ is the empty set, D must in fact be the empty conjunction, which is equivalent to the formula *true*. Thus, for $j = 0$, $T \models D \Rightarrow \varphi$ iff $T \models \varphi$. While $j = 0$ is the case of real interest, the recursive construction Grandjean uses forces us to deal with the case $j > 0$ as well. In this case, the formula $D \Rightarrow \varphi$ contains free variables; these variables are treated as being universally quantified for purposes of determining if $T \models D \Rightarrow \varphi$.

Our algorithm is the natural extension to Grandjean's algorithm for the case of conditional probabilities and for a language with constants. The chief difference is that we begin by considering $T \models D_\mathcal{V} \Rightarrow \varphi$ (where $\mathcal{V}$ is the model fragment on which we are conditioning). Suppose $D_\mathcal{V}$ uses the variables $x_1, \ldots, x_n$, where $n = \nu(\psi)$. We have said that $T \models D_\mathcal{V} \Rightarrow \varphi$ is interpreted as $T \models \forall x_1, \ldots, x_n\ (D_\mathcal{V} \Rightarrow \varphi)$, and this is equivalent to $T \models (\exists x_1, \ldots, x_n\ D_\mathcal{V}) \Rightarrow \varphi$ because φ is closed. Because $\exists x_1, \ldots, x_n\ D_\mathcal{V}$ is in T by definition, this latter assertion is equivalent to $T \models \varphi$, which is what we are really interested in.

Starting from the initial step just outlined, the algorithm then recursively examines smaller and smaller subformulas of φ, while maintaining a description D which keeps track of any new free variables that appear in the current subformula. Of course, D will also extend $D_\mathcal{V}$ and will be consistent with ψ.

We now describe the algorithm in more detail. Without loss of generality, we assume that all negations in φ are pushed in as far as possible, so that only atomic formulas are negated. We also assume that φ does not use the variables $x_1, x_2, x_3, \ldots$. The algorithm proceeds by induction on the structure of the formula, until the base case—an atomic formula or its negation—is reached. The following equivalences form the basis for the recursive procedure:

1. If ξ is of the form ξ' or $\neg\xi'$ for an atomic formula ξ', then $T \models D \Rightarrow \xi$ iff ξ is a conjunct of D.
2. If ξ is of the form $\xi_1 \wedge \xi_2$, then $T \models D \Rightarrow \xi$ iff $T \models D \Rightarrow \xi_1$ and $T \models D \Rightarrow \xi_2$.
3. If ξ is of the form $\xi_1 \vee \xi_2$ then $T \models D \Rightarrow \xi$ iff $T \models D \Rightarrow \xi_1$ or $T \models D \Rightarrow \xi_2$.
4. If ξ is of the form $\exists y\, \xi'$ and D is a complete description over Φ and $\{x_1, \ldots, x_j\}$, then $T \models D \Rightarrow \xi$ iff $T \models D' \Rightarrow \xi'[y/x_{j+1}]$ for some complete description D' over Φ and $\{x_1, \ldots, x_{j+1}\}$ that extends D and is consistent with ψ.
5. If ξ is of the form $\forall y\, \xi'$ and D is a complete description over Φ and $\{x_1, \ldots, x_j\}$, then $T \models D \Rightarrow \xi$ iff $T \models D' \Rightarrow \xi'[y/x_{j+1}]$ for all complete descriptions D'

over Φ and $\{x_1, \ldots, x_{j+1}\}$ that extend D and are consistent with ψ.

The proof that this procedure is correct is based on the following proposition, which can easily be proved using the same techniques as for Proposition 3.25.

PROPOSITION 4.1. *If D is a complete description over Φ and $\mathcal{X}$ and $\xi \in \mathcal{L}(\Phi)$ is a formula all of whose free variables are in $\mathcal{X}$, then either $T \models D \Rightarrow \xi$ or $T \models D \Rightarrow \neg\xi$.*

Proof. We know that T has no finite models. By the Löwenheim–Skolem Theorem [12, p. 141], we can, without loss of generality, restrict attention to countably infinite models of T.

Suppose $\mathcal{X} = \{x_1, x_2, \ldots, x_j\}$ and that $T \not\models D \Rightarrow \xi$. Then there is some countable model $\mathcal{U}$ of T, and j domain elements $\{d_1, \ldots, d_j\}$ in the domain of $\mathcal{U}$, which satisfy $D \wedge \neg\xi$. Consider another model $\mathcal{U}'$ of T, and any $\{d'_1, \ldots, d'_j\}$ in the domain of U' that satisfy D. Because D is a complete description, the substructures over $\{d_1, \ldots, d_j\}$ and $\{d'_1, \ldots, d'_j\}$ are isomorphic. We can use the back and forth construction of Proposition 3.25 to extend this to an isomorphism between $\mathcal{U}$ and $\mathcal{U}'$. But then it follows that $\{d'_1, \ldots, d'_j\}$ must also satisfy $\neg\xi$. Since $\mathcal{U}$ was arbitrary, $T \models D \Rightarrow \neg\xi$. The result follows. ☐

The following result shows that the algorithm above gives a sound and complete procedure for determining whether $T \models D_{\mathcal{V}} \Rightarrow \varphi$.

THEOREM 4.2. *Each of the equivalences in steps* (1)–(5) *above is true.*

Proof. The equivalences for steps (1)–(3) are easy to show, using Proposition 4.1. To prove (4), consider some formula $D \Rightarrow \exists y\, \xi'$, where D is a complete description over $x_1, \ldots, x_j$ and the free variables of ξ are contained in $\{x_1, \ldots, x_j\}$. Let $\mathcal{U}$ be some countable model of T, and let $d_1, \ldots, d_j$ be elements in $\mathcal{U}$ that satisfy D. If $\mathcal{U}$ satisfies $D \Rightarrow \exists y\, \xi'$ then there must exist some other element d_{j+1} that, together with $d_1, \ldots, d_j$, satisfies ξ. Consider the description D' over $x_1, \ldots, x_{j+1}$ that extends D and is satisfied by $d_1, \ldots, d_{j+1}$. Clearly $T \not\models D' \Rightarrow \neg\xi'[y/x_{j+1}]$ because this is false in $\mathcal{U}$. So, by Proposition 4.1, $T \models D' \Rightarrow \xi'[y/x_{j+1}]$ as required.

For the other direction, suppose that $T \models D' \Rightarrow \xi'[y/x_{j+1}]$ for some D' extending D. It follows that $T \models \exists x_{j+1} D' \Rightarrow \exists x_{j+1} \xi'[y/x_{j+1}]$. The result follows from the observation that T contains the extension axiom $\forall x_1, \ldots, x_j (D \Rightarrow \exists x_{j+1} D')$.

The proof for case (5) is similar to that for case (4), and is omitted. ☐

We analyze the complexity of this algorithm in terms of alternating Turing machines (ATMs) [5]. Recall that in an ATM, the nonterminal states are classified into two kinds: universal and existential. Just as with a nondeterministic TM, a nonterminal state may have one or more successors. The terminal states are classified into two kinds: accepting and rejecting. The computation of an ATM forms a tree, where the nodes are instantaneous descriptions (IDs) of the machine's state at various points in the computation, and the children of a node are the possible successor IDs. We recursively define what it means for a node in a computation tree to be an *accepting* node. Leaves are terminal states, and a leaf is accepting just if the machine is in an accepting state in the corresponding ID. A node whose ID is in an existential state is accepting iff at least one of its children is accepting. A node whose ID is in a universal state is accepting iff all of its children are accepting. The entire computation is accepting if the root is an accepting node.

We use several different measures for the complexity of an ATM computation. The time of the computation is the number of steps taken by its longest computation branch. The number of *alternations* of a computation of an ATM is the maximum number of times, over all branches, that the type of state switched (from universal to existential or vice versa). The number of *branches* is simply the number of distinct computation paths. The number of branches is always bounded by an exponential in the computation time, but sometimes we can find tighter bounds.

Grandjean's algorithm, and our variant of it, is easily implemented on an ATM. Each inductive step corresponding to a disjunction or an existential quantifier can be implemented using a sequence of existential guesses. Similarly, each step corresponding to a conjunction or a universal quantifier can be implemented using a sequence of universal guesses. Note that the number of alternations is at most $|\varphi|$. We must analyze the time and branching complexity of this ATM. Given $\psi \wedge \mathcal{V}$, each computation branch of this ATM can be regarded as doing the following. It

(a) constructs a complete description D over the variables $x_1, \ldots, x_{n+k}$ that extends $D_{\mathcal{V}}$ and is consistent with ψ, where $n = \nu(\psi)$ and $k \leq |\varphi|/2$ is the number of variables appearing in φ,

(b) chooses a formula ξ or $\neg\xi$, where ξ is an atomic subformula of φ (with free variables renamed appropriately so that they are included in $\{x_1, \ldots, x_{n+k}\}$), and

(c) checks whether $T \models D \Rightarrow \xi$.

Generating a complete description D requires time $|D|$, and if we construct D by adding conjuncts to $D_{\mathcal{V}}$ then it is necessarily the case that D extends $D_{\mathcal{V}}$. To check whether D is consistent with ψ, we must verify that D does not assert the existence of any new element in any finite atom. Under an appropriate representation of ψ (outlined after Corollary 4.4 below), this check can be done in time $O(|D|2^{|\mathcal{P}|})$. Choosing an atomic subformula ξ of φ can take time $O(|\varphi|)$. Finally, checking whether $T \models D \Rightarrow \xi$ can be accomplished by simply scanning $|D|$. It is easy to see that we can do this without backtracking over $|D|$. Since $|D| > |\xi|$, it can be done in time $O(|D|)$. Combining all these estimates, we conclude that the length of each branch is $O(|D|2^{|\mathcal{P}|} + |\varphi|)$.

Let D be any complete description over Φ and $\mathcal{X}$. Without loss of generality, we assume that each constant in Φ is equal to (at least) one of the variables in $\mathcal{X}$. To fully describe D we must specify, for each predicate R of arity i, which of the i-tuples of variables used in D satisfy R. Thus, the number of choices needed to specify the denotation of R is bounded by $|\mathcal{X}|^\rho$ where ρ is the maximum arity of a predicate in Φ. Therefore, $|D|$ is $O(|\Phi||\mathcal{X}|^\rho)$. In the case of the description D generated by the algorithm, $\mathcal{X}$ is $\{x_1, \ldots, x_n, x_{n+1}, \ldots, x_{n+k}\}$, and $n+k$ is less than $n+|\varphi|$. Thus, the length of such a description D is $O(|\Phi|(n + |\varphi|)^\rho)$.

Using this expression, and our analysis above, we see that the computation time is certainly $O(|\Phi|2^{|\mathcal{P}|}(n + |\varphi|)^\rho)$. In general, the number of branches of the ATM is at most the number of complete descriptions multiplied by the number of atomic formulas in φ. The first of these terms can be exponential in the length of each description. Therefore the number of branches is $O(|\varphi|2^{|\Phi|(n+|\varphi|)^\rho}) = 2^{O(|\Phi|(n+|\varphi|)^\rho)}$. We can, however, get a better bound on the number of branches if all predicates in Φ are unary (i.e., if $\rho = 1$). In this case, ψ already specifies all the properties of the named elements. Therefore, a complete description D is determined when we decide, for each of the at most k variables in D not corresponding to named elements, whether it is equal to a named element and, if not, which atom it satisfies. It follows that there are at most $(2^{|\Phi|}+n)^k$ complete descriptions in this case, and so at most $|\varphi|(2^{|\Phi|}+n)^k$ branches. Since $k \leq |\varphi|/2$, the number of branches is certainly $O((2^{|\Phi|}+n)^{|\varphi|})$ if $\rho = 1$. We summarize this analysis in the following theorem, which forms the basis for almost all of our upper bounds in this section.

THEOREM 4.3. *There exists an alternating Turing machine that takes as input a* *finite vocabulary* Φ*, a model description* $\psi \wedge \mathcal{V}$ *over* Φ*, and a formula* $\varphi \in \mathcal{L}(\Phi)$*, and decides whether* $\Pr^w_\infty(\varphi \mid \psi \wedge \mathcal{V})$ *is* 0 *or* 1*. The machine uses time* $O(|\Phi|2^{|\mathcal{P}|}(\nu(\psi) + |\varphi|)^\rho)$ *and* $O(|\varphi|)$ *alternations, where* ρ *is the maximum arity of predicates in* Φ*. If* $\rho > 1$*, the number of branches is* $2^{O(|\Phi|(\nu(\psi)+|\varphi|)^\rho)}$*. If* $\rho = 1$*, the number of branches is* $O((2^{|\Phi|} + \nu(\psi))^{|\varphi|})$*.*

An alternating Turing machine can be simulated by a deterministic Turing machine which traverses all possible branches of the ATM, while keeping track of the intermediate results necessary to determine whether the ATM accepts or rejects. The time taken by the deterministic simulation is linear in the product of the number of branches of the ATM and the time taken by each branch. The space required is the logarithm of the number of branches plus the space required for each branch. In this case, both these terms are $O(|D| + |\varphi|)$, where D is the description generated by the machine. This allows us to prove the following important corollary.

COROLLARY 4.4. *There exists a deterministic Turing machine that takes as input a finite vocabulary* Φ*, a model description* $\psi \wedge \mathcal{V}$ *over* Φ*, and a formula* $\varphi \in \mathcal{L}(\Phi)$*, and decides whether* $\Pr^w_\infty(\varphi \mid \psi \wedge \mathcal{V})$ *is* 0 *or* 1*. If* $\rho > 1$ *the machine uses time* $2^{O(|\Phi|(\nu(\psi)+|\varphi|)^\rho)}$ *and space* $O(|\Phi|(\nu(\psi) + |\varphi|)^\rho)$*. If* $\rho = 1$ *the machine uses time* $2^{O(|\varphi||\Phi|\log(\nu(\psi)+1))}$ *and space* $O(|\varphi||\Phi|\log(\nu(\psi) + 1))$*.*

4.2. Computing asymptotic conditional probabilities. Our overall goal is to compute $\Pr^w_\infty(\varphi \mid \theta)$ for some $\varphi \in \mathcal{L}(\Phi)$ and $\theta \in \mathcal{L}(\Psi)$. To do this, we enumerate model descriptions over Φ of size $d(\theta) + |\mathcal{C}|$, and check which are consistent with θ. Among those model descriptions that are of maximal degree, we compute the fraction of model descriptions $\psi \wedge \mathcal{V}$ for which $\Pr^w_\infty(\varphi \mid \psi \wedge \mathcal{V})$ is 1.

More precisely, let $\delta_\theta = \Delta^\Psi(\theta)$. Theorem 3.34 tells us that

$$\Pr^w_\infty(\varphi \mid \theta) = \frac{1}{|\mathcal{M}(\mathcal{A}^{\Psi,\delta_\theta}_\theta)|} \sum_{(\psi\wedge\mathcal{V})\in\mathcal{M}(\mathcal{A}^{\Psi,\delta_\theta}_\theta)} \Pr^w_\infty(\varphi \mid \psi \wedge \mathcal{V}).$$

The procedure *Compute-Pr*$_\infty$, described in Fig. 4.1, generates one by one all model descriptions of size $d(\theta) + |\mathcal{C}|$ over Φ. The algorithm keeps track of three things, among the model descriptions considered thus far: (1) the highest degree δ of a model description consistent with θ, (2) the number $count(\theta)$ of model descriptions of degree δ consistent with θ, and (3) among the model descriptions of degree δ consistent with θ, the number $count(\varphi)$ of descriptions such that $\Pr^w_\infty(\varphi \mid \psi \wedge \mathcal{V}) = 1$. Thus, for each model description $\psi \wedge \mathcal{V}$ generated, the algorithm computes $\Delta(\psi)$. If $\Delta(\psi) < \delta$ or $\Pr^w_\infty(\theta \mid \psi \wedge \mathcal{V})$ is 0, then the model description is ignored. Otherwise, if $\Delta(\psi) > \delta$, then the count for lower degrees is irrelevant. In this case, the algorithm erases the previous counts by setting $\delta \leftarrow \Delta(\psi)$, $count(\theta) \leftarrow 1$, and $count(\varphi) \leftarrow \Pr^w_\infty(\varphi \mid \psi \wedge \mathcal{V})$. If $\Delta(\psi) = \delta$, then the algorithm updates $count(\theta)$ and $count(\varphi)$ appropriately.

Examining *Compute-Pr*$_\infty$, we see that its complexity is dominated by two major quantities: the time required to generate all model descriptions, and the time required to compute each 0-1 probability using our variant of Grandjean's algorithm. The complexity of the latter was given in Theorem 4.3 and Corollary 4.4. The following proposition states the length of a model description; the time required to generate all model descriptions is exponential in this length.

PROPOSITION 4.5. *If* $M > |\mathcal{C}|$ *then the length of a model description of size* M *over* Φ *is*

$$O(|\Phi|(2^{|\mathcal{P}|}M)^\rho).$$

Procedure *Compute-Pr*$_\infty(\varphi \mid \theta)$
    $\delta \leftarrow (0,0)$
    For each model description $\psi \wedge \mathcal{V}$ do:
        Compute $\Pr^w_\infty(\theta \mid \psi \wedge \mathcal{V})$ using our variant of Grandjean's algorithm
        If $\Delta(\psi) = \delta$ and $\Pr^w_\infty(\theta \mid \psi \wedge \mathcal{V}) = 1$ then
            $\mathit{count}(\theta) \leftarrow \mathit{count}(\theta) + 1$
            Compute $\Pr^w_\infty(\varphi \mid \psi \wedge \mathcal{V})$ using our variant of Grandjean's algorithm
            $\mathit{count}(\varphi) \leftarrow \mathit{count}(\varphi) + \Pr^w_\infty(\varphi \mid \psi \wedge \mathcal{V})$
        If $\Delta(\psi) > \delta$ and $\Pr^w_\infty(\theta \mid \psi \wedge \mathcal{V}) = 1$ then
            $\delta \leftarrow \Delta(\psi)$
            $\mathit{count}(\theta) \leftarrow 1$
            Compute $\Pr^w_\infty(\varphi \mid \psi \wedge \mathcal{V})$ using our variant of Grandjean's algorithm
            $\mathit{count}(\varphi) \leftarrow \Pr^w_\infty(\varphi \mid \psi \wedge \mathcal{V})$
    If $\delta = (0,0)$ then output "$\Pr^w_\infty(\varphi \mid \theta)$ not well defined"
        otherwise output "$\Pr^w_\infty(\varphi \mid \theta) = \mathit{count}(\varphi)/\mathit{count}(\theta)$".

FIG. 4.1. Compute-$\Pr_\infty$ *for computing asymptotic conditional probabilities.*

Proof. Consider a model description over Φ of size $M = d(\theta) + |\mathcal{C}|$. Such a model description consists of two parts: an atomic description ψ over Ψ and a model fragment $\mathcal{V}$ over Φ which is in $\mathcal{M}(\psi)$. To specify an atomic description ψ, we need to specify the unary properties of the named elements; furthermore, for each atom, we need to say whether it has any elements beyond the named elements (i.e., whether it is active). Using this representation, the size of an atomic description ψ is $O(|\Psi|\nu(\psi) + 2^{|\mathcal{P}|})$. As we have already observed, the length of a complete description D over Φ and $\mathcal{X}$ is $O(|\Phi||\mathcal{X}|^\rho)$. In the case of a description $D_\mathcal{V}$ for $\mathcal{V} \in \mathcal{M}(\psi)$, this is $O(|\Phi|\nu(\psi)^\rho)$. Using $\nu(\psi) \leq 2^{|\mathcal{P}|}M$, we obtain the desired result. ∎

Different variants of this algorithm are the basis for most of the upper bounds in the remainder of this section.

4.3. Finite vocabulary. We now consider the complexity of various problems related to $\Pr^w_\infty(\varphi \mid \theta)$ for a fixed finite vocabulary Φ. The input for such problems is simply φ and θ, and so the input length is the sum of the lengths of φ and θ. Since, for the purposes of this section, we view the vocabulary Φ as fixed (independent of the input), its size and maximum arity can be treated as constants.

We first consider the issue of well-definedness.

THEOREM 4.6. *Fix a finite vocabulary Φ with at least one unary predicate symbol. For $\theta \in \mathcal{L}(\Psi)$, the problem of deciding whether $\Pr^w_\infty(* \mid \theta)$ is well defined is PSPACE-complete. The lower bound holds even if $\theta \in \mathcal{L}^-(\{P\})$.*

Proof. It follows from Lemma 3.30 that $\Pr^w_\infty(* \mid \theta)$ is well defined iff $\alpha^\Psi(\theta) > 0$. This is true iff there is some atomic description $\psi \in \mathcal{A}^\Psi_\theta$ such that $\alpha(\psi) > 0$. This holds iff there exists an atomic description ψ of size $M = d(\theta) + |\mathcal{C}|$ over Ψ and some model fragment $\mathcal{V} \in \mathcal{M}^\Psi(\psi)$ such that $\alpha(\psi) > 0$ and $\Pr^w_\infty(\theta \mid \psi \wedge \mathcal{V}) = 1$. Since we are working within Ψ, we can take $\rho = 1$ and $|\mathcal{P}|$ to be a constant, independent of θ. Thus, the length of a model description $\psi \wedge \mathcal{V}$ as given in Proposition 4.5 is polynomial in $|\theta|$. It is therefore possible to generate model descriptions in PSPACE. Using Corollary 4.4, we can check, in polynomial space, for a model description $\psi \wedge \mathcal{V}$ whether $\Pr^w_\infty(\theta \mid \psi \wedge \mathcal{V})$ is 1. Therefore, the entire procedure can be done in polynomial space.

For the lower bound, we use a reduction from the problem of checking the truth of quantified Boolean formulas (QBF), a problem well known to be PSPACE-complete

[37]. The reduction is similar to that used to show that checking whether a first-order sentence is true in a given finite structure is PSPACE-hard [6]. Given a quantified Boolean formula β, we define a first-order sentence $\xi_\beta \in \mathcal{L}^-(\{P\})$ as follows. The structure of ξ_β is identical to that of β, except that any reference to a propositional variable x, except in the quantifier, is replaced by $P(x)$. For example, if β is $\forall x\, \exists y\, (x \wedge y)$, ξ_β will be $\forall x\, \exists y\, (P(x) \wedge P(y))$. Let θ be $\xi_\beta \wedge \exists x\, P(x) \wedge \exists x\, \neg P(x)$. Clearly, $\Pr^w_\infty(* \mid \theta)$ is well defined exactly if β is true. ☐

In order to compute asymptotic conditional probabilities in this case, we simply use *Compute-Pr*$_\infty$. In fact, since *Compute-Pr*$_\infty$ can also be used to determine well-definedness, we could also have used it to prove the previous theorem.

THEOREM 4.7. *Fix a finite vocabulary* Φ. *For* $\varphi \in \mathcal{L}(\Phi)$ *and* $\theta \in \mathcal{L}(\Psi)$, *the problem of computing* $\Pr^w_\infty(\varphi \mid \theta)$ *is PSPACE-complete. Indeed, deciding if* $\Pr^w_\infty(\varphi \mid \mathit{true}) = 1$ *is PSPACE-hard even if* $\varphi \in \mathcal{L}^-(\{P\})$ *for some unary predicate symbol* P.

Proof. The upper bound is obtained directly from *Compute-Pr*$_\infty$ in Fig. 4.1. The algorithm generates model descriptions one by one. Using the assumption that Φ is fixed and finite, each model description has polynomial length, so that this can be done in PSPACE. Corollary 4.4 implies that, for a fixed finite vocabulary, the 0-1 probabilities for each model description can also be computed in polynomial space. While $\mathit{count}(\theta)$ and $\mathit{count}(\varphi)$ can be exponential (as large as the number of model descriptions), only polynomial space is required for their binary representation. Thus, *Compute-Pr*$_\infty$ works in PSPACE under the assumption of a fixed finite vocabulary.

For the lower bound, we provide a reduction from QBF much like that used in Theorem 4.6. Given a quantified Boolean formula ξ and a unary predicate symbol P, we construct a sentence $\xi_\beta \in \mathcal{L}^-(\{P\})$ just as in the proof of Theorem 4.6. It is easy to see that $\Pr^w_\infty(\xi_\beta \mid \mathit{true}) = 1$ iff β is true. (By the unconditional 0-1 law, $\Pr^w_\infty(\xi_\beta \mid \mathit{true})$ is necessarily either 0 or 1.) ☐

It follows immediately from Theorem 4.7 that we cannot approximate the limit. Indeed, if we fix ϵ with $0 < \epsilon < 1$, the problem of deciding whether $\Pr^w_\infty(\varphi \mid \theta) \in [0, 1-\epsilon]$ is PSPACE-hard even for $\varphi, \theta \in \mathcal{L}^-(\{P\})$. We might hope to prove that for any nontrivial interval $[r_1, r_2]$, it is PSPACE-hard to decide if $\Pr^w_\infty(\varphi \mid \theta) \in [r_1, r_2]$. This stronger lower bound does not hold for the language $\mathcal{L}^-(\{P\})$. Indeed, it follows from Theorem 3.35 that if Φ is any fixed vocabulary then, for $\varphi \in \mathcal{L}(\Phi)$ and $\theta \in \mathcal{L}^-(\Psi)$, $\Pr^w_\infty(\varphi \mid \theta)$ must take one of a finite number of values (the possible values being determined entirely by Φ). So the approximation problem is frequently trivial; in particular, this is the case for any $[r_1, r_2]$ that does not contain one of the possible values. To see that there are only a finite number of values, first note that there is a fixed collection of atoms over Φ. If θ does not use equality, an atomic description can only say, for each atom A over Φ, whether $\exists x\, A(x)$ or $\neg\exists x\, A(x)$ holds. There is also a fixed set of constant symbols to describe. Therefore, there is a fixed set of possible atomic descriptions. Finally, note that the only named elements are the constants, and so there is also a fixed (and finite) set of model fragments. This shows that the set of model descriptions is finite, from which it follows that $\Pr^w_\infty(\varphi \mid \theta)$ takes one of finitely many values fixed by Φ. Thus, in order to have $\Pr^w_\infty(\varphi \mid \theta)$ assume infinitely many values, we must allow equality in the language. Moreover, even with equality in the language, one unary predicate does not suffice. Using Theorem 3.34, it can be shown that two unary predicates are necessary to allow the asymptotic conditional probability to assume infinitely many possible values. As the following result shows, this condition also suffices.

THEOREM 4.8. *Fix a finite vocabulary Φ that contains at least two unary predicates and rational numbers $0 \le r_1 \le r_2 \le 1$ such that $[r_1, r_2] \ne [0, 1]$. For $\varphi, \theta \in \mathcal{L}(\mathcal{P})$, the problem of deciding whether $\Pr_\infty^w(\varphi \mid \theta) \in [r_1, r_2]$ is PSPACE-hard, even given an oracle that tells us whether the limit is well defined.*

Proof. We first show that, for any rational number r with $0 < r < 1$, we can construct φ_r, θ_r such that $\Pr_\infty^w(\varphi_r \mid \theta_r) = r$. Suppose $r = q/p$. We assume, without loss of generality, that $\Phi = \{P, Q\}$. Let θ_r be the sentence

$$\exists^{p-1}x\, P(x) \wedge \left(\exists^{q-1}x\,(P(x) \wedge Q(x)) \vee \exists^{q}x\,(P(x) \wedge Q(x))\right) \wedge \exists^0 x\,(\neg P(x) \wedge \neg Q(x)).$$

That is, no elements satisfy the atom $\neg P \wedge \neg Q$, either q or $q-1$ elements satisfy the atom $P \wedge Q$, and $p-1$ elements satisfy P. Thus, there are exactly two atomic descriptions consistent with θ_r. In one of them, ψ_1, there are $q-1$ elements satisfying $P \wedge Q$ and $p-q$ elements satisfying $P \wedge \neg Q$ (all the remaining elements satisfy $\neg P \wedge Q$). In the other, ψ_2, there are q elements satisfying $P \wedge Q$ and $p-q-1$ elements satisfying $P \wedge \neg Q$. Clearly, the degree of ψ_1 is the same as that of ψ_2, so that neither one dominates. In particular, both define $p-1$ named elements. The number of model fragments for ψ_1 is $\binom{p-1}{q-1} = \frac{(p-1)!}{(q-1)!(p-q)!}$. The number of model fragments for ψ_2 is $\binom{p-1}{q} = \frac{(p-1)!}{q!(p-q-1)!}$. Let φ_r be ψ_1. Clearly

$$\begin{aligned}
\Pr_\infty^w(\varphi_r \mid \theta_r) &= \frac{|\mathcal{M}(\psi_1)|}{|\mathcal{M}(\psi_1)| + |\mathcal{M}(\psi_2)|} \\
&= \frac{(p-1)!/((q-1)!(p-q)!)}{(p-1)!/((q-1)!(p-q)!) + (p-1)!/(q!(p-q-1)!)} \\
&= \frac{q}{q+(p-q)} = \frac{q}{p} = r\ .
\end{aligned}$$

Now, assume we are given $r_1 \le r_2$. We prove the result by reduction from QBF, as in the proof of Theorem 4.6. If $r_1 = 0$ then the result follows immediately from Theorem 4.7. If $0 < r_1 = q/p$, let β be a QBF, and consider $\Pr_\infty^w(\xi_\beta \wedge \varphi_{r_1} \mid \theta_{r_1} \wedge \exists x\, \neg P(x))$. Note that, since $p \ge 2$, θ_{r_1} implies $\exists x\, P(x)$. It is therefore easy to see that this probability is 0 if β is false and $\Pr_\infty^w(\varphi_{r_1} \mid \theta_{r_1}) = r_1$ otherwise. Thus, we can check if β is true by deciding whether $\Pr_\infty^w(\xi_\beta \wedge \varphi_{r_1} \mid \theta_{r_1} \wedge \exists x\, \neg P(x)) \in [r_1, r_2]$. This proves PSPACE-hardness.[9] $\square$

These results show that simply assuming that the vocabulary is fixed and finite is not by itself enough to lead to computationally easy problems. Nevertheless, there is some good news. We observed in a companion paper [23] that if Φ is fixed and finite, and we bound the depth of quantifier nesting, then there exists a linear time algorithm for computing asymptotic probabilities. In general, as we observed in [23], we cannot effectively construct this algorithm, although we know that it exists. As we now show, for the case of conditioning on a unary formula, we can effectively construct this algorithm.

THEOREM 4.9. *Fix $d \ge 0$. For $\varphi \in \mathcal{L}(\Phi)$, $\theta \in \mathcal{L}(\Psi)$ such that $d(\varphi), d(\theta) \le d$, we can effectively construct a linear time algorithm that decides if $\Pr_\infty^w(\varphi \mid \theta)$ is well defined and computes it if it is.*

[9] In this construction, it is important to note that although φ_{r_1} and θ_{r_1} can be long sentences, their length depends only on r_1, which is treated as being fixed. Therefore, the constructed asymptotic probability expression does have length polynomial in $|\beta|$. This is also the case in similar constructions later in the paper.

Proof. The proof of the general theorem in [23] shows that if there is a bound d on the quantification depth of formulas and a finite vocabulary, then there is a finite set Σ_d of formulas such that every formula ξ of depth at most d is equivalent to a formula in Σ_d. Moreover, we can construct an algorithm that, given such a formula ξ, will in linear time find some formula equivalent to ξ in Σ_d. (We say "some" rather than "the," because it is necessary for the algorithm's constructibility that there will generally be several formulas equivalent to ξ in Σ_d.) Given this, the problem reduces to constructing a lookup table for the asymptotic conditional probabilities for all formulas in Σ_d. In general, there is no effective technique for constructing this table. However, if we allow conditioning only on unary formulas, it follows from Theorem 4.7 that there is. The result now follows. ◻

4.4. Infinite vocabulary—restricted cases. In the next two sections we consider an infinite vocabulary Ω. As discussed in §2.3, there are at least two distinct interpretations for asymptotic conditional probabilities in the case of an infinite vocabulary. One interpretation of "infinite vocabulary" views Ω as a potential or background vocabulary, so that every problem instance includes as part of its input the actual finite subvocabulary that is of interest. So, although this subvocabulary is finite, there is no bound on its possible size. The alternative is to interpret infinite vocabularies more literally, using the limit process explained in §2.3. In the case of the random-worlds method, Proposition 2.1 shows that both interpretations give the same result. Thus, it is immediate that all complexity results we prove with respect to one interpretation immediately hold for the other. As we are postponing the discussion of random structures to §4.6, we present the earlier results with respect to the second, less cumbersome, interpretation.

As before, we are interested in computing the complexity of the same three problems: deciding whether the asymptotic probability is well defined, computing it, and approximating it. As we mentioned earlier, the complexity is quite sensitive to a number of factors. One factor, already observed in the unconditional case [4], [20], is whether there is a bound on the arity of the predicates in Ω. Without such a bound, the problem is complete for the class #TA(EXP,LIN). Unlike the unconditional case, however, simply putting a bound on the arity of the predicates in Ω is not enough to improve the complexity (unless the bound is 1); we also need to restrict the use of equality, so that it cannot appear in the right-hand side of the conditional. Roughly speaking, with equality, we can use the named elements to play the same role as the predicates of unbounded arity. In this section, we consider what happens if we in fact restrict the language so that either (1) Ω has no predicate of arity ≥ 2, or (2) there is a bound (which may be greater than 1) on the arity of the predicates in Ω, but we never condition on formulas that use equality. As we now show, these two cases turn out to be quite similar. In particular, the same complexity results hold.

Throughout this section, we take Ω to be a fixed infinite vocabulary such that all predicate symbols in Ω have arity less than some fixed bound ρ. Let $\mathcal{Q}$ be the set of all unary predicate symbols in Ω, let $\mathcal{D}$ be the set of all constant symbols in Ω, and let $\Upsilon = \mathcal{Q} \cup \mathcal{D}$.

We start with the problem of deciding whether the asymptotic probability is well defined. Since well-definedness depends only on the right-hand side of the conditional, which we already assume is restricted to mentioning only unary predicates, its complexity is independent of the bound ρ.

The following theorem, due to Lewis [30], is the key to proving the lower bound for well-definedness (and for some of the other results in this section as well).

THEOREM 4.10. [30] *The problem of deciding whether a sentence $\xi \in \mathcal{L}^-(\mathcal{Q})$ is satisfiable is NEXPTIME-complete. Moreover, the lower bound holds even for formulas ξ of depth 2.*

Lewis proves this as follows: for any nondeterministic Turing machine $\mathbf{M}$ that runs in exponential time and any input w, he constructs a sentence $\xi \in \mathcal{L}^-(\mathcal{Q})$ of quantifier depth 2 and whose length is polynomial in the size of $\mathbf{M}$ and w, such that ξ is satisfiable iff there is an accepting computation of $\mathbf{M}$ on w.

Our first use of Lewis's result is to show that determining well-definedness is NEXPTIME-complete; this result does not require the assumptions that we are making throughout the rest of this section.

THEOREM 4.11. *For $\theta \in \mathcal{L}(\Upsilon)$, the problem of deciding if $\Pr^w_\infty(* \mid \theta)$ is well defined is NEXPTIME-complete. The NEXPTIME lower bound holds even for $\theta \in \mathcal{L}^-(\mathcal{Q})$ where $d(\theta) \leq 2$.*

Proof. For the upper bound, we proceed much as in Theorem 4.6. Let $\Psi = \Upsilon_\theta$ and let $\mathcal{C} = \mathcal{D}_\theta$. We know that $\Pr^w_\infty(* \mid \theta)$ is well defined iff there exists an atomic description ψ of size $M = d(\theta) + |\mathcal{C}|$ over Ψ and some model fragment $\mathcal{V} \in \mathcal{M}^\Psi(\psi)$ such that $\alpha(\psi) > 0$ and $\Pr^w_\infty(\theta \mid \psi \wedge \mathcal{V}) = 1$. Since all the predicates in Ψ have arity 1, it follows from Proposition 4.5 that the size of a model description $\psi \wedge \mathcal{V}$ over Ψ is $O(|\Psi| 2^{|\mathcal{P}|} M)$. Since $|\Psi| < |\theta|$, this implies that model descriptions have exponential length, and can be generated by a nondeterministic exponential time Turing machine. Because we can assume that $\rho = 1$ here when applying Corollary 4.4, we can also deduce that we can check whether $\Pr^w_\infty(\theta \mid \psi \wedge \mathcal{V})$ is 0 or 1 using a deterministic Turing machine in time $2^{O(|\theta||\Psi| \log(\nu(\psi)+1))}$. Since $|\Psi| \leq |\theta|$, and $\nu(\psi)$ is at most exponential in $|\theta|$, it follows that we can decide if $\Pr^w_\infty(\theta \mid \psi \wedge \mathcal{V}) = 1$ in deterministic time exponential in $|\theta|$. Thus, to check if $\Pr^w_\infty(* \mid \theta)$ is well defined we nondeterministically guess a model description $\psi \wedge \mathcal{V}$ of the right type, and check that $\alpha(\psi) > 0$ and that $\Pr^w_\infty(\theta \mid \psi \wedge \mathcal{V}) = 1$. The entire procedure can be executed in nondeterministic exponential time.

For the lower bound, observe that if a formula ξ in $\mathcal{L}^-(\Phi)$ is satisfied in some model with domain $\{1, \ldots, N\}$ then it is satisfiable in some model of every domain size larger than N. Therefore, $\xi \in \mathcal{L}^-(\mathcal{Q})$ is satisfiable if and only if the limit $\Pr^w_\infty(* \mid \xi)$ is well defined. The result now follows from Theorem 4.10. ☐

We next consider the problem of computing the asymptotic probability $\Pr^w_\infty(\varphi \mid \theta)$, given that it is well defined. We show that this problem is #EXP-complete. Recall that #P (see [38]) is the class of integer functions computable as the number of accepting computations of a nondeterministic polynomial-time Turing machine. More precisely, a function $f : \{0,1\}^* \to \mathbb{N}$ is said to be in #P if there is a nondeterministic Turing machine $\mathbf{M}$ such that for any w, the number of accepting paths of $\mathbf{M}$ on input w is $f(w)$. The class #EXP is the exponential time analogue.

The function we are interested in is $\Pr^w_\infty(\varphi \mid \theta)$, which is not integer valued. Nevertheless, we want to show that it is in #EXP. In the spirit of similar definitions for #P (see, for example, [39] and [34]) and NP (e.g., [17]) we extend the definition of #EXP to apply also to non-integer-valued functions.

DEFINITION 4.12. *An arbitrary function f is said to be* #EXP-easy *if there exists an integer-valued function g in #EXP and a polynomial-time-computable function h such that for all x, $f(x) = h(g(x))$. (In particular, we allow h to involve divisions, so that $f(x)$ may be a rational function.) A function f is* #EXP-hard *if, for every #EXP-easy function g, there exist polynomial-time functions h_1 and h_2 such that, for* *all x, $g(x) = h_2(f(h_1(x)))$.*[10] *A function f is* #EXP-complete *if it is #EXP-easy and #EXP-hard.* □

We can similarly define analogues of these definitions for the class #P.

We now show that for an infinite arity-bounded vocabulary in which equality is not used, or for any unary vocabulary, the problem of computing the asymptotic conditional probability is #EXP-complete. We start with the upper bound.

THEOREM 4.13. *If either* (a) $\varphi, \theta \in \mathcal{L}(\Upsilon)$ *or* (b) $\varphi \in \mathcal{L}(\Omega)$ *and* $\theta \in \mathcal{L}^{-}(\Upsilon)$*, then computing* $\Pr^w_\infty(\varphi \mid \theta)$ *is #EXP-easy.*

Proof. Let $\Phi = \Omega_{\varphi\wedge\theta}$, let $\Psi = \Upsilon_{\varphi\wedge\theta}$, and let $\mathcal{P}$ and $\mathcal{C}$ be the appropriate subsets of Ψ. Let $\delta_\theta = \Delta^\Psi(\theta)$. Recall from the proof of Theorem 4.7 that we would like to generate the model descriptions $\psi \wedge \mathcal{V}$ of degree δ_θ, consider the ones for which $\Pr^w_\infty(\theta \mid \psi \wedge \mathcal{V}) = 1$, and compute the fraction of those for which $\Pr^w_\infty(\varphi \mid \psi \wedge \mathcal{V})$. More precisely, consider the set of model descriptions of size $M = d(\varphi \wedge \theta) + |\mathcal{C}|$. For a degree δ, let $\mathit{count}^\delta(\theta)$ denote the number of those model descriptions for which $\Pr^w_\infty(\theta \mid \psi \wedge \mathcal{V}) = 1$. Similarly, let $\mathit{count}^\delta(\varphi)$ denote the number for which $\Pr^w_\infty(\varphi \wedge \theta \mid \psi \wedge \mathcal{V}) = 1$. We are interested in the value of the fraction $\mathit{count}^{\delta_\theta}(\varphi)/\mathit{count}^{\delta_\theta}(\theta)$.

We want to show that we can nondeterministically generate model descriptions $\psi \wedge \mathcal{V}$, and check in deterministic exponential time whether $\Pr^w_\infty(\theta \mid \psi \wedge \mathcal{V})$ (or, similarly, $\Pr^w_\infty(\varphi \wedge \theta \mid \psi \wedge \mathcal{V})$) is 0 or 1. We begin by showing the second part: that the 0-1 probabilities can be computed in deterministic exponential time. There are two cases to consider. In case (a), φ and θ are both unary, allowing us to assume that $\rho = 1$ for the purposes of Corollary 4.4. In this case, the 0-1 computations can be done in time $2^{O(|\varphi\wedge\theta||\Psi|\log(\nu(\psi)+1))}$, where $\Psi = \Upsilon_{\varphi\wedge\theta}$. As in Theorem 4.11, $|\Psi| \le |\varphi \wedge \theta|$ and $\nu(\psi)$ is at most exponential in $|\theta|$, allowing us to carry out the computation in deterministic exponential time. In case (b), $\theta \in \mathcal{L}^{-}(\Upsilon)$, and therefore the only named elements are the constants. In this case, the 0-1 probabilities can be computed in deterministic time $2^{O(|\Phi|(\nu(\psi)+|\varphi\wedge\theta|)^\rho)}$, where $\Phi = \Omega_{\varphi\wedge\theta}$. However, as we have just discussed, $\nu(\psi) < |\varphi \wedge \theta|$, implying that the computation can be completed in exponential time.

Having shown how the 0-1 probabilities can be computed, it remains only to generate model descriptions in the appropriate way. However, we do not want to consider all model descriptions, because we must count only those model descriptions of degree δ_θ. The problem is that we do not know δ_θ in advance. We will therefore construct a nondeterministic exponential time Turing machine $\mathbf{M}$ such that the number of accepting paths of $\mathbf{M}$ encodes, for each degree δ, both $\mathit{count}^\delta(\varphi)$ and $\mathit{count}^\delta(\theta)$. We need to do the encoding in such a way as to be able to isolate the counts for δ_θ when we finally know its value. This is done as follows.

Let ψ be an atomic description ψ over Ψ of size M. Recall that the degree $\Delta(\psi)$ is a pair $(\alpha(\psi), \nu(\psi))$ such that $\alpha(\psi) \le 2^{|\mathcal{P}|}$ and $\nu(\psi) \le 2^{|\mathcal{P}|}M$. Thus, there are at most $E = 2^{2|\mathcal{P}|}M$ possible degrees. Number the degrees in increasing order: $\delta_1, \ldots, \delta_E$. We want it to be the case that the number of accepting paths of $\mathbf{M}$ written in binary has the form

$$p_{10}\cdots p_{1m}q_{10}\cdots q_{1m}\cdots p_{E0}\cdots p_{Em}q_{E0}\cdots q_{Em},$$

where $p_{i0}\ldots p_{im}$ is the binary representation of $\mathit{count}^{\delta_i}(\varphi)$ and $q_{i0}\ldots q_{im}$ is the binary representation of $\mathit{count}^{\delta_i}(\theta)$. We choose m to be sufficiently large so that there is no

[10] Notice that we need the function h_2 as well as h_1. For example, if g is an integer-valued function and f always returns a rational value between 0 and 1, as is the case for us, then there is no function h_1 such that $g(x) = f(h_1(x))$.

overlap between the different sections of the output. The largest possible value of an expression of the form $count^{\delta_i}(\theta)$ is the maximum number of model descriptions of degree δ_i over Φ. This is clearly less than the overall number of model descriptions, which we computed in §4.2.

The machine $\mathbf{M}$ proceeds as follows. Let m be the smallest integer such that 2^m is more than the number of possible model descriptions, which, by Proposition 4.5, is $2^{O(|\Phi|(2^{|\mathcal{P}|}M)^\rho)}$. Note that m is exponential and that $\mathbf{M}$ can easily compute m from Φ. $\mathbf{M}$ then nondeterministically chooses a degree δ_i, for $i = 1, \ldots, E$. It then executes $E - i$ phases, in each of which it nondeterministically branches $2m$ times. This has the effect of giving this branch a weight of $2^{2m(E-i)}$. It then nondeterministically chooses whether to compute $p_{i0} \ldots p_{im}$ or $q_{i0} \ldots q_{im}$. If the former, it again branches m times, separating the results for $count^{\delta_i}(\varphi)$ from those for $count^{\delta_i}(\theta)$. In either case, it now nondeterministically generates all model descriptions $\psi \wedge \mathcal{V}$ over Φ. It ignores those for which $\Delta(\psi) \neq \delta_i$. For the remaining model descriptions $\psi \wedge \mathcal{V}$, it computes $\Pr^w_\infty(\varphi \wedge \theta \mid \psi \wedge \mathcal{V})$ in the first case, and $\Pr^w_\infty(\theta \mid \psi \wedge \mathcal{V})$ in the latter. This is done in exponential time, using the same technique as in Theorem 4.11. The machine accepts precisely when this probability is 1.

This procedure is executable in nondeterministic exponential time, and results in the appropriate number of accepting paths. It is now easy to compute $\Pr^w_\infty(\varphi \mid \theta)$ by finding the largest degree δ for which $count^\delta(\theta) > 0$, and dividing $count^\delta(\varphi)$ by $count^\delta(\theta)$. ☐

We now want to prove the matching lower bound. As in Theorem 4.11, we make use of Lewis's NEXPTIME-completeness result. As there, this allows us to prove the result even for $\varphi, \theta \in \mathcal{L}^-(\mathcal{Q})$ of quantifier depth 2. A straightforward modification of Lewis's proof shows that, given w and a nondeterministic exponential time Turing machine $\mathbf{M}$, we can construct a depth-2 formula $\xi \in \mathcal{L}^-(\mathcal{Q})$ such that the number of simplified atomic descriptions over $\mathcal{P}_\xi$ consistent with ξ is exactly the number of accepting computations of $\mathbf{M}$ on w. This allows us to prove the following theorem.

THEOREM 4.14. *Given $\xi \in \mathcal{L}^-(\mathcal{Q})$, counting the number of simplified atomic descriptions over $\mathcal{P}_\xi$ consistent with ξ is #EXP-complete. The lower bound holds even for formulas ξ of depth* 2.

This theorem forms the basis for our own hardness result.

THEOREM 4.15. *Given $\varphi, \theta \in \mathcal{L}^-(\mathcal{Q})$ of depth at least* 2*, the problem of computing $\Pr^w_\infty(\varphi \mid \theta)$ is #EXP-hard, even given an oracle for deciding whether the limit exists.*

Proof. Given $\varphi \in \mathcal{L}^-(\mathcal{Q})$, we reduce the problem of counting the number of simplified atomic descriptions over $\mathcal{P}_\varphi$ consistent with φ to that of computing an appropriate asymptotic probability. Recall that, for the language $\mathcal{L}^-(\mathcal{Q})$, model descriptions are equivalent to simplified atomic descriptions. Therefore, computing an asymptotic conditional probability for this language reduces to counting simplified atomic descriptions of maximal degree. Thus, the major difficulty we need to overcome here is the converse of the difficulty that arose in the upper bound. We now want to count *all* simplified atomic descriptions consistent with φ, while using the asymptotic conditional probability in the most obvious way would only let us count those of maximum degree. For example, the two atomic descriptions whose characteristics are represented in Fig. 4.2 have different degrees; the first one will thus be ignored by a computation of asymptotic conditional probabilities.

Let $\mathcal{P}$ be $\mathcal{P}_\varphi = \{P_1, \ldots, P_k\}$, and let Q be a new unary predicate not in $\mathcal{P}$. We let $A_1, \ldots, A_K$ for $K = 2^k$ be all the atoms over $\mathcal{P}$, and let $A'_1, \ldots, A'_{2K}$ be all the atoms over $\mathcal{P}' = \mathcal{P} \cup \{Q\}$, such that $A'_i = A_i \wedge Q$ and $A'_{K+i} = A_i \wedge \neg Q$ for $i = 1, \ldots, K$.

A_1	A_2	A_3	A_4
$*$	0	$*$	0
$*$	$*$	0	$*$

FIG. 4.2. *Two atomic descriptions with different degrees.*

	A_1	A_2	A_3	A_4
$\wedge Q$:	$*$	0	$*$	0
$\wedge\neg Q$:	0	$*$	0	$*$
$\wedge Q$:	$*$	$*$	0	$*$
$\wedge\neg Q$:	0	0	$*$	0

FIG. 4.3. *Two maximal atomic descriptions.*

We define θ' as follows:

$$\theta' =_{\rm def} \forall x, y \left(\left(Q(x) \wedge \bigwedge_{i=1}^{k} (P_i(x) \Leftrightarrow P_i(y)) \right) \Rightarrow Q(y) \right) .$$

The sentence θ' guarantees that the predicate Q is "constant" on the atoms defined by $\mathcal{P}$. That is, if A_i is an atom over $\mathcal{P}$, it is not possible to have $\exists x\,(A_i(x) \wedge Q(x))$ as well as $\exists x\,(A_i(x) \wedge \neg Q(x))$. Therefore, if ψ is a simplified atomic description over $\mathcal{P}'$ which is consistent with θ', then, for each $i \leq K$, at most one of the atoms A'_i and A'_{K+i} can be active, while the other is necessarily empty. It follows that $\alpha(\psi) \leq K$. Since there are clearly atomic descriptions of activity count K consistent with θ', the atomic descriptions of maximal degree are precisely those for which $\alpha(\psi) = K$. Moreover, if $\alpha(\psi) = K$, then A'_i is active iff A'_{K+i} is not. Two atomic descriptions of maximal degree are represented in Fig. 4.3. Thus, for each set $I \subseteq \{1, \ldots, K\}$, there is precisely one simplified atomic description ψ consistent with θ' of activity count K where A'_i is active in ψ iff $i \in I$. Therefore, there are exactly 2^K simplified atomic descriptions ψ over $\mathcal{P}'$ consistent with θ' for which $\alpha(\psi) = K$.

Let $\theta = \theta' \wedge \exists x\, Q(x)$. Notice that all simplified atomic descriptions ψ with $\alpha(\psi) = K$ that are consistent with θ' are also consistent with θ, except for the one where no atom in $A'_1, \ldots, A'_K$ is active. Thus, $|\mathcal{A}_\theta^{\mathcal{P}',K}| = 2^K - 1$. For the purposes of this proof, we call a simplified atomic description ψ over $\mathcal{P}'$ consistent with θ for which $\alpha(\psi) = K$ a *maximal atomic description.* Notice that there is an obvious one-to-one correspondence between consistent simplified atomic descriptions over $\mathcal{P}$ and maximal atomic descriptions over $\mathcal{P}'$. A maximal atomic description where A'_i is active iff $i \in I$ (and A'_{K+i} is active for $i \notin I$) corresponds to the simplified atomic description over $\mathcal{P}$ where A_i is active iff $i \in I$. (For example, the two consistent simplified atomic descriptions over $\{P_1, P_2\}$ in Fig. 4.2 correspond to the two maximal atomic descriptions over $\{P_1, P_2, Q\}$ in Fig. 4.3.) In fact, the reason we consider θ rather than θ' is precisely because there is no consistent simplified atomic description over $\mathcal{P}$ which corresponds to the maximal atomic description where no atom in $A'_1, \ldots, A'_K$ is active (since there is no consistent atomic description over $\mathcal{P}$ where none of $A_1, \ldots, A_K$ are active). Thus, we have overcome the hurdle discussed above.

We now define φ_Q; intuitively, φ_Q is φ restricted to elements that satisfy Q. Formally, we define ξ_Q for any formula ξ by induction on the structure of the formula:

- $\xi_Q = \xi$ for any atomic formula ξ,
- $(\neg\xi)_Q = \neg\xi_Q$,
- $(\xi \wedge \xi')_Q = \xi_Q \wedge \xi'_Q$,
- $(\forall y\, \xi(y))_Q = \forall y\, (Q(y) \Rightarrow \xi_Q(y))$.

Note that the size of φ_Q is linear in the size of φ. The one-to-one mapping discussed above from simplified atomic descriptions to maximal atomic descriptions gives us a one-to-one mapping from simplified atomic descriptions over $\mathcal{P}$ consistent with φ to maximal atomic descriptions consistent with $\varphi_Q \wedge \exists x\, Q(x)$. This is true because a model satisfies φ_Q iff the same model restricted to elements satisfying Q satisfies φ. Thus, the number of model descriptions over $\mathcal{P}$ consistent with φ is precisely $|\mathcal{A}^{\mathcal{P}',K}_{\varphi_Q \wedge \theta}|$.

From Corollary 3.36, it follows that

$$\Pr^w_\infty(\varphi_Q \mid \theta) = \frac{|\mathcal{A}^{\mathcal{P}',K}_{\varphi_Q \wedge \theta}|}{|\mathcal{A}^{\mathcal{P}',K}_{\theta}|} = \frac{|\mathcal{A}^{\mathcal{P}}_{\varphi}|}{2^K - 1}.$$

Thus, the number of simplified atomic descriptions over $\mathcal{P}$ consistent with φ is $(2^K - 1)\Pr^w_\infty(\varphi_Q \mid \theta)$. This proves the lower bound. □

As in Theorem 4.8, we can also show that any nontrivial approximation of the asymptotic probability is hard, even if we restrict to sentences of depth 2.

THEOREM 4.16. *Fix rational numbers $0 \le r_1 \le r_2 \le 1$ such that $[r_1, r_2] \neq [0, 1]$. For $\varphi, \theta \in \mathcal{L}^-(\mathcal{Q})$ of depth at least 2, the problem of deciding whether $\Pr^w_\infty(\varphi \mid \theta) \in [r_1, r_2]$ is both NEXPTIME-hard and co-NEXPTIME-hard, even given an oracle for deciding whether the limit exists.*

Proof. Let us begin with the case where $r_1 = 0$ and $r_2 < 1$. Consider any $\varphi \in \mathcal{L}^-(\mathcal{Q})$ of depth at least 2, and assume without loss of generality that $\mathcal{P} = \mathcal{P}_\varphi = \{P_1, \ldots, P_k\}$. Choose $Q \notin \mathcal{P}$, let $\mathcal{P}' = \mathcal{P} \cup \{Q\}$, and let ξ be $\forall x (P_1(x) \wedge \ldots \wedge P_k(x) \wedge Q(x))$. We consider $\Pr^w_\infty(\varphi \mid \varphi \vee \xi)$. Clearly $\varphi \vee \xi$ is satisfiable, so that this asymptotic probability is well defined. If φ is unsatisfiable, then $\Pr^w_\infty(\varphi \mid \varphi \vee \xi) = 0$. On the other hand, if φ is satisfiable, then $\alpha^{\mathcal{P}}(\varphi) = j > 0$ for some j. It is easy to see that $\alpha^{\mathcal{P}'}(\varphi) = \alpha^{\mathcal{P}'}(\varphi \vee \xi) = 2j$. Moreover, φ and $\varphi \vee \xi$ are consistent with precisely the same simplified atomic descriptions with $2j$ active atoms. This is true since $\alpha^{\mathcal{P}'}(\xi) = 1 < 2j$. It follows that if φ is satisfiable, then $\Pr^w_\infty(\varphi \mid \varphi \vee \xi) = 1$.

Thus, we have that $\Pr^w_\infty(\varphi \mid \varphi \vee \xi)$ is either 0 or 1, depending on whether or not φ is satisfiable. Thus, $\Pr^w_\infty(\neg\varphi \mid \varphi \vee \xi)$ is in $[r_1, r_2]$ iff φ is satisfiable; similarly, $\Pr^w_\infty(\neg\varphi \mid \neg\varphi \vee \xi)$ is in $[r_1, r_2]$ iff φ is valid. By Theorem 4.10, it follows that this approximation problem is both NEXPTIME-hard and co-NEXPTIME-hard.

If $r_1 = q/p > 0$, we construct sentences φ_{r_1} and θ_{r_1} of depth 2 in $\mathcal{L}^-(\mathcal{Q})$ such that $\Pr^w_\infty(\varphi_{r_1} \mid \theta_{r_1}) = r_1$.[11] Choose $\ell = \lceil \log p \rceil$, and let $\mathcal{P}'' = \{Q_1, \ldots, Q_\ell\}$ be a set of predicates such that $\mathcal{P}'' \cap \mathcal{P}' = \emptyset$. Let $A_1, \ldots, A_{2^\ell}$ be the set of atoms over $\mathcal{P}''$. We define θ_{r_1} to be

$$\exists^1 x\, (A_1(x) \vee A_2(x) \vee \ldots \vee A_p(x)).$$

Similarly, φ_{r_1} is defined as

$$\exists^1 x\, (A_1(x) \vee A_2(x) \vee \ldots \vee A_q(x)).$$

[11] The sentences constructed in Theorem 4.8 for the same purpose will not serve our purpose in this theorem, since they used unbounded quantifier depth.

Recall from §3.1 that the construct "$\exists^1 x$" can be defined in terms of a formula of quantifier depth 2. There are exactly p atomic descriptions of size 2 of maximal degree consistent with θ_{r_1}; each has one element in one of the atoms $A_1, \ldots, A_p$ and no elements in the rest of the atoms among $A_1, \ldots, A_p$, with all the remaining atoms (those among $A_{p+1}, \ldots, A_{2^\ell}$) being active. Among these atomic descriptions, q are also consistent with φ_{r_1}. Therefore, $\Pr^w_\infty(\varphi_{r_1} \mid \theta_{r_1}) = q/p$. Since the predicates occurring in $\varphi_{r_1}, \theta_{r_1}$ are disjoint from $\mathcal{P}'$, it follows that

$$\Pr^w_\infty(\varphi \wedge \varphi_{r_1} \mid (\varphi \vee \xi) \wedge \theta_{r_1}) = \Pr^w_\infty(\varphi \mid \varphi \vee \xi) \cdot \Pr^w_\infty(\varphi_{r_1} \mid \theta_{r_1}) = \Pr^w_\infty(\varphi \mid \varphi \vee \xi) \cdot r_1.$$

This is equal to r_1 (and hence is in $[r_1, r_2]$) if and only if φ is satisfiable, and is 0 other- wise. □

The lower bounds in this section all hold provided we consider formulas whose quantification depth is at least 2. Can we do better if we restrict to formulas of quantification depth at most 1? As is suggested by Table 1.1, we can. The complexities typically drop by an exponential factor. For example, checking well-definedness becomes NP-complete rather than NEXPTIME-complete. For the problem of computing probabilities for formulas with quantification depth 1, we know that the problem is in PSPACE, and is (at least) #P-hard. Finally, the problem of approximating probabilities is hard for both NP and co-NP. A detailed analysis of these results can be found in [27]; some related work for a propositional language has been done by Roth [35].

4.5. Infinite vocabulary—the general case. In §4.4 we investigated the complexity of asymptotic conditional probabilities when the (infinite) vocabulary satisfies certain restrictions. As we now show, the results there were tight in the sense that the restrictions cannot be weakened. We examine the complexity of the general case, in which the vocabulary is infinite with no bound on predicates' arities and/or in which equality can be used.

The problem of checking if $\Pr^w_\infty(\varphi \mid \theta)$ is well defined is still NEXPTIME-complete. Theorem 4.11 (which had no restrictions) still applies. However, the complexity of the other problems we consider does increase. It can be best described in terms of the complexity class TA(EXP,LIN)—the class of problems that can be solved by an exponential time ATM with a linear number of alternations. The class TA(EXP,LIN) also arises in the study of unconditional probabilities where there is no bound on the arity of the predicates. Grandjean [20] proved a TA(EXP,LIN) upper bound for computing whether the unconditional probability is 0 or 1 in this case, and Immerman [4] proved a matching lower bound. Of course, Grandjean's result can be viewed as a corollary of Theorem 4.3. Immerman's result, which has not, to the best of our knowledge, appeared in print, is a corollary of Theorem 4.18 which we prove in this section.

To capture the complexity of computing the asymptotic probability in the general case, we use a counting class #TA(EXP,LIN) that corresponds to TA(EXP,LIN). To define this class, we restrict attention to the class of ATMs whose initial states are existential. Given such an ATM $\mathbf{M}$, we define an *initial existential path* in the computation tree of $\mathbf{M}$ on input w to be a path in this tree, starting at the initial state, such that every node on the path corresponds to an existential state except for the last node, which corresponds to a universal or an accepting state. That is, an initial existential path is a maximal path that starts at the root of the tree and contains only existential nodes except for the last node in the path. We say that an integer-valued function $f : \{0,1\}^* \to I\!N$ is in #TA(EXP,LIN) if there is a machine $\mathbf{M}$ in the class TA(EXP,LIN) such that, for all w, $f(w)$ is the number of existential paths in the computation tree of $\mathbf{M}$ on input w whose last node is accepting (recall that we defined a notion of "accepting" for any node in the tree in §4.1). We extend the definition of #TA(EXP,LIN) to apply to non-integer-valued functions and define *#TA(EXP,LIN)-easy* just as we did before with #P and #EXP in §4.4.

We start with the upper bound.

THEOREM 4.17. *For* $\varphi \in \mathcal{L}(\Omega)$ *and* $\theta \in \mathcal{L}(\Upsilon)$*, the function* $\Pr_\infty^w(\varphi \mid \theta)$ *is in #TA(EXP,LIN).*

Proof. Let $\Phi = \Omega_{\varphi\wedge\theta}$, let $\Psi = \Upsilon_{\varphi\wedge\theta}$, and let ρ be the maximum arity of a predicate in Φ. The proof proceeds precisely as in Theorem 4.13. We compute, for each degree δ, the values $\mathit{count}^\delta(\theta)$ and $\mathit{count}^\delta(\varphi)$. This is done by nondeterministically generating model descriptions $\psi \wedge \mathcal{V}$ over Φ, branching according to the degree of ψ, and computing $\Pr_\infty^w(\varphi \wedge \theta \mid \psi \wedge \mathcal{V})$ and $\Pr_\infty^w(\theta \mid \psi \wedge \mathcal{V})$ using a TA(EXP,LIN) Turing machine.

To see that this is possible, recall from Proposition 4.5 that the length of a model description over Φ is $O(|\Phi|(2^{|\mathcal{P}|}M)^\rho)$. This is exponential in $|\Phi|$ and ρ, both of which are at most $|\varphi \wedge \theta|$. Therefore, it is possible to guess a model description in exponential time. Similarly, as we saw in the proof of Theorem 4.13, only exponentially many nondeterministic guesses are required to separate the output so that counts corresponding to different degrees do not overlap. These guesses form the initial nondeterministic stage of our TA(EXP,LIN) Turing machine. Note that it is necessary to construct the rest of the Turing machine so that a universal state always follows this initial stage, so that each guess corresponds exactly to one initial existential path; however, this is easy to arrange.

For each model description $\psi \wedge \mathcal{V}$ so generated, we compute $\Pr_\infty^w(\theta \mid \psi \wedge \mathcal{V})$ or $\Pr_\infty^w(\varphi \wedge \theta \mid \psi \wedge \mathcal{V})$ as appropriate, accepting if the conditional probability is 1. It follows immediately from Theorem 4.3 and the fact that there can only be exponentially many named elements in any model description we generate that this computation is in TA(EXP,LIN). Thus, the problem of computing $\Pr_\infty^w(\varphi \mid \theta)$ is in #TA(EXP,LIN). ◻

We now want to prove the matching lower bound. Moreover, we would like to show that the restrictions from §4.4 cannot be weakened. Recall from Theorem 4.13 that the #EXP upper bound held under one of two conditions: either (a) φ and θ are both unary, or (b) the vocabulary is arity-bounded and θ does not use equality. To show that (a) is tight, we show that the #TA(EXP,LIN) lower bound holds even if we allow φ and θ to use only binary predicates and equality. (The use of equality is necessary, since without it we know from (b) that the problem is #EXP-easy.) To show that (b) is tight, we show that the lower bound holds for a non-arity-bounded vocabulary, but without allowing equality in θ. Neither lower bound requires the use of constants.

The proof of the lower bounds is lengthy, but can be simplified somewhat by some assumptions about the construction of the TA(EXP,LIN) machines we consider. The main idea is that the existential "guesses" being made in the initial phase should be clearly distinguished from the rest of the computation. To achieve this, we assume that the Turing machine has an additional *guess tape*, and the initial phase of every computation consists of nondeterministically generating a *guess string* γ which is written on the new tape. The machine then proceeds with a standard alternating computation, but with the possibility of reading the bits on the guess tape.

More precisely, from now on we make the following assumptions about an ATM $\mathbf{M}$. Consider any increasing functions $T(n)$ and $A(n)$ (in essence, these correspond to the time complexity and number of alternations), and let w be an input of size n. We assume:

- $\mathbf{M}$ has two tapes and two heads (one for each tape). Both tapes are one-way infinite to the right.
- The first tape is a work tape, which initially contains only the string w.
- $\mathbf{M}$ has an initial nondeterministic phase, during which its only action is to nondeterministically generate a string γ of zeros and ones, and write this string on the second tape (the guess tape). The string γ is always of length $T(n)$. Moreover, at the end of this phase, the work tape is as in the initial configuration, the guess tape contains only γ, the heads are at the beginning of their respective tapes, and the machine is in a distinguished universal state s_0.
- After the initial phase, the guess tape is never changed.
- After the initial phase, $\mathbf{M}$ takes at most $T(n)$ steps on each branch of its computation tree, and makes exactly $A(n) - 1$ alternations before entering a terminal (accepting or rejecting) state.
- The state before entering a terminal state is always an existential state (i.e., $A(n)$ is odd).

Let $\mathbf{M}'$ be any (unrestricted) TA(T,A) machine that "computes" an integer function f. It is easy to construct some $\mathbf{M}$ satisfying the restrictions above that also computes f. The machine $\mathbf{M}$ first generates the guess string γ, and then simulates $\mathbf{M}'$. At each nondeterministic branching point in the initial existential phase of $\mathbf{M}'$, $\mathbf{M}$ uses the next bit of the string γ to dictate which choice to take. Observe that this phase is deterministic (given γ), and can thus be folded into the following universal phase. (Deterministic steps can be viewed as universal steps with a single successor.) If not all the bits in γ are used, $\mathbf{M}$ continues the execution of $\mathbf{M}'$, but checks in parallel that the unused bits of γ are all 0's. If not, $\mathbf{M}$ rejects. It is easy to see that on any input w, $\mathbf{M}$ has the same number of accepting paths as $\mathbf{M}'$, and therefore accepts the same function f. Moreover, $\mathbf{M}$ has the same number of alternations as $\mathbf{M}'$, and at most a constant factor blowup in the running time.[12] This shows that it will be sufficient to prove our hardness results for the class #TA(EXP,LIN) by considering only those machines that satisfy these restrictions. For the remainder of this section we will therefore assume that all ATMs are of this type.

Let $\mathbf{M}$ be such an ATM and let w be an input of size n. We would like to encode the computation of $\mathbf{M}$ on w using a pair of formulas φ_w, θ_w. (Of course, these formulas depend on $\mathbf{M}$ as well, but we suppress this dependence.) Our first theorem shows how to encode part of this computation: given some appropriate string γ of length $T(n)$, we construct formulas that encode the computation of $\mathbf{M}$ immediately following the initial phase of guessing γ. More precisely, we say that $\mathbf{M}$ *accepts w given γ* if, on input w, the initial existential path during which $\mathbf{M}$ writes γ on the guess tape leads to an accepting node. We construct formulas $\varphi_{w,\gamma}$ and $\theta_{w,\gamma}$ such that $\Pr_\infty^w(\varphi_{w,\gamma} \mid \theta_{w,\gamma})$ is either 0 or 1, and is equal to 1 iff $\mathbf{M}$ accepts w given γ.

We do not immediately want to specify the process of guessing γ, so our initial construction will not commit to this. For a predicate R, let $\varphi[R]$ be a formula that uses the predicate R. Let ξ be another formula that has the same number of free variables as the arity of R. Then $\varphi[\xi]$ is the formula where every occurrence of R is replaced with the formula ξ, with an appropriate substitution of the arguments of R

[12]For ease of presentation, we can and will (somewhat inaccurately, but harmlessly) ignore this constant factor and say that the time complexity of $\mathbf{M}$ is, in fact, $T(n)$.

for the free variables in ξ.

THEOREM 4.18. *Let* $\mathbf{M}$ *be a TA(T,A) machine as above, where* $T(n) = 2^{t(n)}$ *for some polynomial* $t(n)$ *and* $A(n) = O(n)$. *Let* w *be an input string of length* n, *and* $\gamma \in \{0,1\}^{T(n)}$ *be a guess string.*

(a) *For a unary predicate* R, *there exist formulas* $\varphi_w[R], \xi_\gamma \in \mathcal{L}(\Omega)$ *and* $\theta_w \in \mathcal{L}(\Upsilon)$ *such that* $\Pr^w_\infty(\varphi_w[\xi_\gamma] \mid \theta_w) = 1$ *iff* $\mathbf{M}$ *accepts* w *given* γ *and is* 0 *otherwise. Moreover,* φ_w *uses only predicates with arity* 2 *or less.*

(b) *For a binary predicate* R, *there exist formulas* $\varphi'_w[R], \xi'_\gamma \in \mathcal{L}(\Omega)$ *such that* $\Pr^w_\infty(\varphi'_w[\xi'_\gamma] \mid true) = 1$ *iff* $\mathbf{M}$ *accepts* w *given* γ *and is* 0 *otherwise.*

The formulas $\varphi_w[R]$, θ_w, *and* $\varphi'_w[R]$ *are independent of* γ, *and their length is polynomial in the representation of* $\mathbf{M}$ *and* w. *Moreover, none of the formulas constructed use any constant symbols.*

Proof. Let Γ be the tape alphabet of $\mathbf{M}$ and let S be the set of states of $\mathbf{M}$. We will identify an instantaneous description (ID) of length ℓ of $\mathbf{M}$ with a string Σ^ℓ for $\Sigma = \Sigma_W \times \Sigma_G$, where Σ_W is $\Gamma \cup (\Gamma \times S)$ and Σ_G is $(\{0,1\} \cup (\{0,1\} \times \{h\}))$. We think of the Σ_W component of the ith element in a string as describing the contents of the ith location in the work tape and also, if the tape head is at location i, the state of the Turing machine. The Σ_G component describes the contents of the ith location in the guess tape (whose alphabet is $\{0,1\}$) and whether the guess tape's head is positioned there. Of course, we consider only strings in which exactly one element in $\Gamma \times S$ appears in the first component and exactly one element in $\{0,1\} \times \{h\}$ appears in the second component. Since $\mathbf{M}$ halts within $T(n)$ steps (not counting the guessing process, which we treat separately), we need only deal with IDs of length at most $T(n)$. Without loss of generality, assume all IDs have length exactly $T(n)$. (If necessary, we can pad shorter IDs with blanks.)

In both parts of the theorem, IDs are encoded using the properties of domain elements. In both cases, the vocabulary contains predicates whose truth value with respect to certain combinations of domain elements represent IDs. The only difference between parts (a) and (b) is in the precise encoding used. We begin by showing the encoding for part (a).

In part (a), we use the sentence θ_w to define $T(n)$ named elements. This is possible since θ_w is allowed to use equality. Each ID of the machine will be represented using a single domain element. The properties of the ID will be encoded using the relations between the domain element representing it and the named elements. More precisely, assume that the vocabulary has $t(n)$ unary predicates $P_1, \ldots, P_{t(n)}$, and one additional unary predicate P^*. The domain is divided into two parts: the elements satisfying P^* are the named elements used in the process of encoding IDs, while the elements satisfying $\neg P^*$ are used to actually represent IDs. The formula θ_w asserts (using equality) that each of the atoms A over $\{P^*, P_1, \ldots, P_{t(n)}\}$ in which P^* (as opposed to $\neg P^*$) is one of the conjuncts contains precisely one element:

$$\forall x, y \left(\left(P^*(x) \land P^*(y) \land \bigwedge_{i=1}^{t(n)} (P_i(x) \Leftrightarrow P_i(y)) \right) \Rightarrow x = y \right).$$

Note that θ_w has polynomial length and is independent of γ.

We can view an atom A over $\{P^*, P_1, \ldots, P_{t(n)}\}$ in which P^* is one of the conjuncts as encoding a number between 0 and $T(n) - 1$, written in binary: if A contains P_j rather than $\neg P_j$, then the jth bit of the encoded number is 1; otherwise it is 0. (Recall that $T(n)$, the running time of $\mathbf{M}$, is $2^{t(n)}$.) In the following, we let A_i, for

$i = 0, \ldots, T(n) - 1$, denote the atom corresponding to the number i according to this scheme. Let e_i be the unique element in the atom A_i for $i = 0, \ldots, T(n) - 1$. When representing an ID using a domain element d (where $\neg P^*(d)$), the relation between d and e_i is used to represent the ith coordinate in the ID represented by d. Assume that the vocabulary has a binary predicate R_σ for each $\sigma \in \Sigma$. Roughly speaking, we say that the domain element d represents the ID $\sigma_0 \ldots \sigma_{T(n)-1}$ if $R_{\sigma_i}(d, e_i)$ holds for $i = 0, \ldots, T(n) - 1$. More precisely, we say that d *represents* $\sigma_0 \ldots \sigma_{T(n)-1}$ if

$$\neg P^*(d) \wedge \bigwedge_{i=0}^{T(n)-1} \forall y \left(A_i(y) \Rightarrow \left(R_{\sigma_i}(d, y) \wedge \bigwedge_{\sigma' \in \Sigma - \{\sigma_i\}} \neg R_{\sigma'}(d, y) \right) \right).$$

Note that not every domain element d such that $\neg P^*(d)$ holds encodes a valid ID. However, the question of which ID, if any, is encoded by a domain element d depends only on the relations between d and the finite set of elements $e_0, \ldots, e_{T(n)-1}$. This implies that, with asymptotic probability 1, every ID will be encoded by some domain element. More precisely, let $ID(x) = \sigma_0 \ldots \sigma_{T(n)-1}$ be a formula which is true if x denotes an element that represents $\sigma_0 \ldots \sigma_{T(n)-1}$. (It should be clear that such a formula is indeed expressible in our language.) Then for each ID $\sigma_0 \ldots \sigma_{T(n)-1}$ we have

$$\Pr_\infty^w(\exists x\, (ID(x) = \sigma_0 \ldots \sigma_{T(n)-1}) \mid \theta_w) = 1.$$

For part (b) of the theorem, we must represent IDs in a different way because we are not allowed to condition on formulas that use equality. Therefore, we cannot create an exponential number of named elements using a polynomial-sized formula. The encoding we use in this case uses two domain elements per ID rather than one. We now assume that the vocabulary Ω contains a $t(n)$-ary predicate R_σ for each symbol $\sigma \in \Sigma$. Note that this uses the assumption that there is no bound on the arity of predicates in Ω. For $i = 0, \ldots, T(n) - 1$, let $b^i_{t(n)} \ldots b^i_1$ be the binary encoding of i. We say that the pair (d_0, d_1) of domain elements *represents the ID* $\sigma_0 \ldots \sigma_{T(n)-1}$ if

$$d_0 \neq d_1 \wedge \bigwedge_{i=0}^{T(n)-1} \left(R_{\sigma_i}(d_{b^i_1}, \ldots, d_{b^i_{t(n)}}) \wedge \bigwedge_{\sigma' \in \Sigma - \{\sigma_i\}} \neg R_{\sigma'}(d_{b^i_1}, \ldots, d_{b^i_{t(n)}}) \right).$$

Again, we can define a formula in our language $ID(x_0, x_1) = \sigma_0 \ldots \sigma_{T(n)-1}$ which is true if x_0, x_1 denote a pair of elements that represent $\sigma_0 \ldots \sigma_{T(n)-1}$. As before, observe that for each ID $\sigma_0 \ldots \sigma_{T(n)-1}$ we have

$$\Pr_\infty^w(\exists x_0, x_1\, (ID(x_0, x_1) = \sigma_0 \ldots \sigma_{T(n)-1}) \mid \mathit{true}) = 1.$$

In both case (a) and case (b), we can construct formulas polynomial in the size of $\mathbf{M}$ and w that assert certain properties. For example, in case (a), $Rep(x)$ is true of a domain element d if and only if d encodes an ID. In this case, $Rep(x)$ is the formula

$$\neg P^*(x) \wedge \forall y \left(P^*(y) \Rightarrow \dot{\bigvee}_{\sigma \in \Sigma} R_\sigma(x, y) \right) \wedge$$
$$\exists! y\, (P^*(y) \wedge \bigvee_{\sigma \in ((\Gamma \times S) \times \Sigma_G)} R_\sigma(x, y)) \wedge \exists! y\, (P^*(y) \wedge \bigvee_{\sigma \in (\Sigma_W \times (\{0,1\} \times \{h\}))} R_\sigma(x, y))$$

where $\dot{\bigvee}$ is an abbreviation whose meaning is that precisely one of its disjuncts is true.

In case (b), $Rep(x_0, x_1)$ is true of a pair (d_0, d_1) if and only if it encodes an ID. The construction is similar. For instance, the conjunct of $Rep(x_0, x_1)$ asserting that each tape position has a uniquely defined content is

$$x_0 \neq x_1 \wedge \forall z_1, \ldots, z_{t(n)} \left(\left(\bigwedge_{i=1}^{t(n)} (z_i = x_0 \vee z_i = x_1) \right) \Rightarrow \dot{\bigvee_{\sigma \in \Sigma}} R_\sigma(z_1, \ldots, z_{t(n)}) \right).$$

Except for this assertion, the construction for the two cases is completely parallel given the encoding of IDs. We will therefore restrict the remainder of the discussion to case (a). Other relevant properties of an ID that we can formulate are:

- *Acc*(x) (resp., *Univ*(x), *Exis*(x)) is true of a domain element d if and only if d encodes an ID and the state in *ID*(d) is an accepting state (resp., a universal state, an existential state).
- *Step*(x, x') is true of elements d and d' if and only if both d and d' encode IDs and *ID*(d') can follow from *ID*(d) in one step of $\mathbf{M}$.
- *Comp*(x, x') is true of elements d and d' if and only if both d and d' encode IDs, and *ID*(d') is the final ID in a maximal nonalternating path starting at *ID*(d) in the computation tree of $\mathbf{M}$, and the length of this path is at most $T(n)$. A maximal nonalternating path is either a path all of whose states are existential except for the last one (which must be universal or accepting), or a path all of whose states are universal except for the last one. We can construct *Comp* using a divide and conquer argument, so that its length is polynomial in $t(n)$.

We remark that *Acc*, *Step*, etc. are not new predicate symbols in the language. Rather, they are complex formulas described in terms of the basic predicates R_σ. We omit details of their construction here; these can be found in [20].

It remains only to describe the formula that encodes the initial configuration of $\mathbf{M}$ on input w. Since we are interested in the behavior of $\mathbf{M}$ given a particular guess string γ, we begin by encoding the computation of $\mathbf{M}$ after the initial nondeterministic phase; that is, after the string γ is already written on the guess tape and the rest of the machine is back in its original state. We now construct the formula *Init*$[R](x)$ that describes the initial configuration. This formula takes R as a parameter, and has the form *Init*$'(x) \wedge R(x)$. The formulas substituted for $R(x)$ will correspond (in a way discussed below) to possible guesses γ.

We begin by considering case (a). We assume the existence of an additional binary predicate B_0. It is easy to write a polynomial-length formula *Init*$'(x)$ which is true of a domain element d if and only if d represents an ID where: (a) the state is the distinguished state s_0 entered after the nondeterministic guessing phase, (b) the work tape contains only w, (c) the heads are at the beginning of their respective tapes, and (d) for all i, the ith location of the guess tape contains 0 iff $B_0(d, e_i)$. Here e_i is, as before, the unique element in atom A_i. Note that the last constraint can be represented polynomially using the formula

$$\forall y \left(P^*(y) \Rightarrow \left(B_0(x, y) \Leftrightarrow \bigvee_{\sigma \in \Sigma_W \times \{0, (0,h)\}} R_\sigma(x, y) \right) \right).$$

We also want to find a formula ξ_γ that can constrain B_0 to reflect the guess γ. This formula, which serves as a possible instantiation for R, does not have to be of polynomial size. We define it as follows, where for convenience, we use B_1 as an abbreviation for $\neg B_0$:

$$\xi_\gamma(x) =_{\text{def}} \bigwedge_{i=0}^{T(n)-1} \forall y\ (A_i(y) \Rightarrow B_{\gamma_i}(x,y))\ . \tag{4.1}$$

Note that this is of exponential length.

In case (b), the relation of the guess string γ to the initial configuration is essentially the same modulo the modifications necessary due to the different representation of IDs. We only sketch the construction. As in case (a), we add a predicate B'_0, but in this case of arity $t(n)$. Again, the predicate B'_0 represents the locations of the 0's in the guess tape following the initial nondeterministic phase. The specification of the denotation of this predicate is done using an exponential-sized formula ξ'_γ, as follows (again taking B'_1 to be an abbreviation for $\neg B'_0$):

$$\xi'_\gamma(x_0,x_1) =_{\text{def}} B'_{\gamma_0}(x_0,\ldots,x_0,x_0) \wedge B'_{\gamma_1}(x_0,\ldots,x_0,x_1) \wedge \ldots \wedge B'_{\gamma_{T(n)-1}}(x_1,\ldots,x_1,x_1).$$

Using these formulas, we can now write a formula expressing the assertion that $\mathbf{M}$ accepts w given γ. In writing these formulas, we make use of the assumptions made about $\mathbf{M}$ (that it is initially in the state immediately following the initial guessing phase, that all computation paths make exactly $A(n)$ alternations, and so on). The formula $\varphi_w[R]$ has the following form:

$$\begin{aligned}\exists x_1\,(\mathit{Init}[R](x_1) \wedge \forall x_2\,(\mathit{Comp}(x_1,x_2) \Rightarrow \exists x_3\,(\mathit{Comp}(x_2,x_3) \wedge \forall x_4\,(\mathit{Comp}(x_3,x_4) \Rightarrow \ldots \\ \exists x_{A(n)}\,(\mathit{Comp}(x_{A(n)-1},x_{A(n)}) \wedge \mathit{Acc}(x_{A(n)}))\ldots)))).\end{aligned}$$

It is clear from the construction that $\varphi_w[R]$ does not depend on γ and that its length is polynomial in the representations of $\mathbf{M}$ and w.

Now suppose $\mathcal{W}$ is a world satisfying θ_w in which every possible ID is represented by at least one domain element. (As we remarked above, a random world has this property with asymptotic probability 1.) Then it is straightforward to verify that $\varphi_w[\xi_\gamma]$ is true in $\mathcal{W}$ iff $\mathbf{M}$ accepts w. Therefore $\Pr^w_\infty(\varphi_w[\xi_\gamma] \mid \theta_w) = 1$ iff $\mathbf{M}$ accepts w given γ and 0 otherwise. Similarly, in case (b), we have shown the construction of analogous formulas $\varphi'_w[R]$, for a binary predicate R, and ξ'_γ such that $\Pr^w_\infty(\varphi'_w[\xi'_\gamma] \mid \mathit{true}) = 1$ iff $\mathbf{M}$ accepts w given γ, and is 0 otherwise. ☐

We can now use the above theorem in order to prove the #TA(EXP,LIN) lower bound.

THEOREM 4.19. *For $\varphi \in \mathcal{L}(\Omega)$ and $\theta \in \mathcal{L}(\Upsilon)$, computing $\Pr^w_\infty(\varphi \mid \theta)$ is #TA(EXP,LIN)-hard. The lower bound holds even if φ, θ do not mention constant symbols and either* (a) *φ uses no predicate of arity > 2, or* (b) *θ uses no equality.*

Proof. Let $\mathbf{M}$ be a TA(EXP,LIN) Turing machine of the restricted type discussed earlier, and let w be an input of size n. We would like to construct formulas φ, θ such that from $\Pr^w_\infty(\varphi \mid \theta)$ we can derive the number of accepting computations of $\mathbf{M}$ on w. The number of accepting initial existential paths of such a Turing machine is precisely the number of guess strings γ such that $\mathbf{M}$ accepts w given γ. In Theorem 4.18, we showed how to encode the computation of such a machine $\mathbf{M}$ on input w given a nondeterministic guess γ. We now show how to force an asymptotic conditional probability to count guess strings appropriately.

As in Theorem 4.18, let $T(n) = 2^{t(n)}$, and let $\mathcal{P}' = \{P'_1, \ldots, P'_{t(n)}\}$ be new unary predicates not used in the construction of Theorem 4.18. As before, we can view an atom A' over $\mathcal{P}'$ as representing a number in the range $0, \ldots, T(n)-1$: if A contains P'_j, then the jth bit of the encoded number is 1; otherwise it is 0. Again, let A'_i, for

$i = 0, \ldots, T(n) - 1$, denote the atom corresponding to the number i according to this scheme. We can view a simplified atomic description ψ over $\mathcal{P}'$ as representing the string $\gamma = \gamma_0 \ldots \gamma_{T(n)-1}$ such that γ_i is 1 if ψ contains the conjunct $\exists z\, A'_i(z)$, and 0 if ψ contains its negation. Under this representation, for every string γ of length $T(n)$, there is a unique simplified atomic description over $\mathcal{P}'$ that represents it; we denote this atomic description ψ_γ. Note that ψ_γ is not necessarily a consistent atomic description, since the atomic description where all atoms are empty also denotes a legal string—that string where all bits are 0.

While it is possible to reduce the problem of counting accepting guess strings to that of counting simplified atomic descriptions, this is not enough. After all, we have already seen that computing asymptotic conditional probabilities ignores all atomic descriptions that are not of maximal degree. We deal with this problem as in Theorem 4.15. Let Q be a new unary predicate, and let θ' be, as in Theorem 4.15, the sentence

$$\forall x, y \left(\left(Q(x) \wedge \bigwedge_{j=1}^{t(n)} (P'_j(x) \Leftrightarrow P'_j(y)) \right) \Rightarrow Q(y) \right).$$

Observe that here we use θ' rather than the formula θ of Theorem 4.15, since we also want to count the "inconsistent" atomic description where all atoms are empty. Recall that, assuming θ', each simplified atomic description ψ_γ over $\mathcal{P}'$ corresponds precisely to a single maximal atomic description ψ'_γ over $\mathcal{P}' \cup \{Q\}$. We reduce the problem of counting accepting guess strings to that of counting maximal atomic descriptions over $P' \cup \{Q\}$.

We now consider cases (a) and (b) separately, beginning with the former. Fix a guess string γ. In Theorem 4.18, we constructed formulas $\varphi_w[R], \xi_\gamma \in \mathcal{L}(\Omega)$ and $\theta_w \in \mathcal{L}(\Upsilon)$ such that $\Pr^w_\infty(\varphi_w[\xi_\gamma] \mid \theta_w) = 1$ iff $\mathbf{M}$ accepts w given γ, and is 0 otherwise. Recall that the formula $\xi_\gamma(x)$ (see equation (4.1)) sets the ith guess bit to be γ_i by forcing the appropriate one of $B_0(x, e_i)$ and $B_1(x, e_i)$ to hold, where e_i is the unique element in the atom A_i. In Theorem 4.18, this was done directly by reference to the bits γ_i. Now, we want to derive the correct bit values from ψ_γ, which tells us that the ith bit is 1 iff $\exists z\, A'_i(z)$. The following formula ξ has precisely the desired property:

$$\xi(x) =_{\text{def}} \forall y \left(P^*(y) \Rightarrow \left(B_1(x, y) \Leftrightarrow \exists z \left(Q(z) \wedge \bigwedge_{j=1}^{t(n)} (P_j(y) \Leftrightarrow P'_j(z)) \right) \right) \right).$$

Clearly, $\psi'_\gamma \models \xi \Leftrightarrow \xi_\gamma$.

Similarly, for case (b), the formula ξ' is:

$$\xi'(x_0, x_1) =_{\text{def}} \forall y_1, \ldots, y_{t(n)} \left(\left(\bigwedge_{j=1}^{t(n)} (y_j = x_0 \vee y_j = x_1) \right) \Rightarrow \right.$$
$$\left. \left(B'_1(y_1, \ldots, y_{t(n)}) \Leftrightarrow \exists z \left(Q(z) \wedge \bigwedge_{j=1}^{t(n)} (y_j = x_1 \Leftrightarrow P'_j(z)) \right) \right) \right).$$

As in part (a), $\psi'_\gamma \models \xi' \Leftrightarrow \xi'_\gamma$.

Now, for case (a), we want to compute the asymptotic conditional probability $\Pr^w_\infty(\varphi[\xi] \mid \theta_w \wedge \theta')$. In doing this computation, we will use the observation (whose straightforward proof we leave to the reader) that if the symbols that appear in θ_2 are disjoint from those that appear in φ_1 and θ_1, then $\Pr^w_\infty(\varphi_1 \mid \theta_1 \wedge \theta_2) = \Pr^w_\infty(\varphi_1 \mid \theta_1)$. Using this observation and the fact that all maximal atomic descriptions over $\mathcal{P}' \cup \{Q\}$ are equally likely given $\theta_w \wedge \theta'$, by straightforward probabilistic reasoning we obtain:

$$\begin{aligned}\Pr^w_\infty(\varphi_w[\xi] \mid \theta_w \wedge \theta') &= \sum_{\psi'_\gamma} \Pr^w_\infty(\varphi_w[\xi] \mid \theta_w \wedge \theta' \wedge \psi'_\gamma) \cdot \Pr^w_\infty(\psi'_\gamma \mid \theta_w \wedge \theta') \\ &= \frac{1}{2^{T(n)}} \sum_{\psi'_\gamma} \Pr^w_\infty(\varphi_w[\xi] \mid \theta_w \wedge \theta' \wedge \psi'_\gamma).\end{aligned}$$

We observed before that ξ is equivalent to ξ_γ in worlds satisfying ψ'_γ, and therefore

$$\Pr^w_\infty(\varphi_w[\xi] \mid \theta_w \wedge \theta' \wedge \psi'_\gamma) = \Pr^w_\infty(\varphi_w[\xi_\gamma] \mid \theta_w \wedge \theta' \wedge \psi'_\gamma) = \Pr^w_\infty(\varphi_w[\xi_\gamma] \mid \theta_w),$$

where the second equality follows from the observation that none of the vocabulary symbols in ψ'_γ or θ' appear anywhere in $\varphi_w[\xi_\gamma]$ or in θ_w. In Theorem 4.18, we proved that $\Pr^w_\infty(\varphi_w[\xi_\gamma] \mid \theta_w)$ is equal to 1 if the ATM accepts w given γ and 0 if not. We therefore obtain that

$$\Pr^w_\infty(\varphi_w[\xi] \mid \theta_w \wedge \theta') = \frac{f(w)}{2^{T(n)}}.$$

Since both $\varphi_w[\xi]$ and $\theta_w \wedge \theta'$ are polynomial in the size of the representation of $\mathbf{M}$ and in $n = |w|$, this concludes the proof for part (a). The completion of the proof for part (b) is essentially identical. $\square$

It remains only to investigate the problem of approximating $\Pr^w_\infty(\varphi \mid \theta)$ for this language.

THEOREM 4.20. *Fix rational numbers $0 \leq r_1 \leq r_2 \leq 1$ such that $[r_1, r_2] \neq [0,1]$. For $\varphi, \theta \in \mathcal{L}(\Omega)$, the problem of deciding whether $\Pr^w_\infty(\varphi \mid \theta) \in [r_1, r_2]$ is TA(EXP,LIN)-hard, even given an oracle for deciding whether the limit exists.*

Proof. For the case of $r_1 = 0$ and $r_2 < 1$, the result is an easy corollary of Theorem 4.18. We can generalize this to the case of $r_1 > 0$, using precisely the same technique as in Theorem 4.16. $\square$

4.6. Complexity for random structures. So far in this section, we have investigated the complexity of various problems relating to the asymptotic conditional probability using the random-worlds method. We now deal with the same issues for the case of random structures. It turns out that most of our results for random worlds carry through to random structures for trivial reasons.

First, consider the issue of well-definedness. By Proposition 2.3, well-definedness is equivalent for random worlds and random structures. Therefore, all of the results obtained for random worlds carry through unchanged for random structures.

For computing or approximating the limit, Theorem 3.37 allows us to restrict attention to unary vocabularies and unary sentences φ and θ. In particular, there is no need to duplicate the results in §4.5. For the remainder of this section, we analyze the complexity of computing $\Pr^{s,\Psi}_\infty(\varphi \mid \theta)$ for $\varphi, \theta \in \mathcal{L}(\Psi)$. As before, we can assume that $\mathcal{A}^\Psi_{\varphi\wedge\theta} \subseteq \mathcal{A}^\Psi_\theta$.

The computational approach is essentially the same as that for random worlds. However, as we showed in §3.5, rather than partitioning θ into model descriptions, we can make use of the assumption that the vocabulary is unary and instead partition it into atomic descriptions ψ. That is, for $a = \alpha^\Psi(\theta)$,

$$\Pr^{s,\Psi}_\infty(\varphi \mid \theta) = \frac{1}{|\mathcal{A}^{\Psi,a}_\theta|} \cdot \sum_{\psi \in \mathcal{A}^{\Psi,a}} \Pr^w_\infty(\varphi \mid \psi) \ \left(= \frac{|\mathcal{A}^{\Psi,a}_{\varphi\wedge\theta}|}{|\mathcal{A}^{\Psi,a}_\theta|} \right).$$

As for random worlds, we begin with the problem of computing 0-1 probabilities. In §4.1, we showed how to extend Grandjean's algorithm to compute $\Pr^w_\infty(\varphi \mid \psi \wedge \mathcal{V})$. Fix a unary vocabulary Ψ, and suppose that $\psi \wedge \mathcal{V}$ is a model description over Ψ, with $n = \nu(\psi)$. Recall from Proposition 3.21 that $\Pr^w_\infty(\varphi \mid \psi \wedge \mathcal{V}) = \Pr^w_\infty(\varphi \mid \psi \wedge \exists x_1, \ldots, x_n\, D_{\mathcal{V}})$. However, in the unary case it is easy to see that $\psi \wedge \exists x_1, \ldots, x_n\, D_{\mathcal{V}}$ is equivalent to ψ. This is because the only nontrivial properties of the named elements given by $\mathcal{V}$ is which atom each of them satisfies and the equality relations between the constants, and this information is already present in the atomic description ψ.

Therefore, we conclude that $\Pr^w_\infty(\varphi \mid \psi)$ is either 0 or 1, because this is so for $\Pr^w_\infty(\varphi \mid \psi \wedge \mathcal{V})$. Now recall that if $\psi \in \mathcal{A}^\Psi_\varphi$ then ψ implies φ. In this case, clearly $\Pr^w_\infty(\varphi \mid \psi) = \Pr^{s,\Psi}_\infty(\varphi \mid \psi) = 1$. Similarly, if $\psi \notin \mathcal{A}^\Psi_\varphi$, then ψ is inconsistent with φ and $\Pr^w_\infty(\varphi \mid \psi) = \Pr^{s,\Psi}_\infty(\varphi \mid \psi) = 0$. So it follows that we can continue to use Grandjean's algorithm, as described in §4.1, to compute $\Pr^{s,\Psi}_\infty(\varphi \mid \psi)$.

THEOREM 4.21. *There exists an alternating Turing machine that takes as input a finite unary vocabulary Ψ, an atomic description ψ over Ψ, and a formula $\varphi \in \mathcal{L}(\Psi)$, and decides whether $\Pr^{s,\Psi}_\infty(\varphi \mid \psi)$ is 0 or 1. The machine uses time $O(|\Psi|2^{|\mathcal{P}|}(\nu(\psi) + |\varphi|))$ and has at most $O((2^{|\Psi|} + \nu(\psi))^{|\varphi|})$ branches and $O(|\varphi|)$ alternations.*

As before, we can simulate the ATM deterministically.

COROLLARY 4.22. *There exists a deterministic Turing machine that takes as input a finite unary vocabulary Ψ, an atomic description ψ over Ψ, and a formula $\varphi \in \mathcal{L}(\Psi)$, and decides whether $\Pr^{w,\Psi}_\infty(\varphi \mid \psi)$ is 0 or 1. The machine uses time $2^{O(|\varphi||\Psi|\log(\nu(\psi)+1))}$ and space $O(|\varphi||\Psi|\log(\nu(\psi) + 1))$.*

We now analyze the complexity of computing $\Pr^{s,\Psi}_\infty(\varphi \mid \theta)$. We begin with the case of a fixed finite vocabulary Ψ.

THEOREM 4.23. *Fix a finite unary vocabulary Ψ with at least one predicate symbol. For $\varphi, \theta \in \mathcal{L}(\Psi)$, the problem of computing $\Pr^{s,\Psi}_\infty(\varphi \mid \theta)$ is PSPACE-complete. Moreover, deciding if $\Pr^w_\infty(\varphi \mid \mathit{true}) = 1$ is PSPACE-hard, even if $\varphi \in \mathcal{L}^-(\{P\})$.*

Proof. By Corollaries 3.41 and 3.42, if $\varphi, \theta \in \mathcal{L}^-(\{P\})$ and $P \in \Psi$, then $\Pr^w_\infty(\varphi \mid \theta) = \Pr^{s,\{P\}}_\infty(\varphi \mid \theta) = \Pr^{s,\Psi}_\infty(\varphi \mid \theta)$. Thus, the lower bound follows immediately from Theorem 4.7.

For the upper bound, we follow the same general procedure of *Compute-Pr*$_\infty$: generating all atomic descriptions of size $d(\theta) + |\mathcal{C}|$, and computing $\Pr^{s,\Psi}_\infty(\varphi \mid \theta)$. The only difference is that, rather than counting only model descriptions of the highest degree $\Delta(\psi) = (\alpha(\psi), \nu(\psi))$, we count all atomic descriptions of the highest activity count $\alpha(\psi)$. Clearly, since there are fewer atomic descriptions than model descriptions, and an atomic description has a shorter description length than a model description, the complexity of the resulting algorithm can only be lower than the corresponding complexity for random worlds. The algorithm for random structures is therefore also in PSPACE. □

Just as Theorem 4.7, Theorem 4.23 shows that even approximating the limit is hard. That is, for a fixed ϵ with $0 < \epsilon < 1$, the problem of deciding whether $\Pr^{s,\Psi}_\infty(\varphi \mid \theta) \in [0, 1 - \epsilon]$ is PSPACE-hard even for $\varphi, \theta \in \mathcal{L}^-(\{P\})$. As for the case of random worlds, this lower bound cannot be generalized to arbitrary intervals $[r_1, r_2]$ unless we allow equality. In particular, for any fixed finite language, there is a fixed number of atomic descriptions of size 1, where this number depends only on the language. Therefore, there are only finitely many values that the probability $\Pr^{s,\Psi}_\infty(\varphi \mid \theta)$ can take for $\varphi, \theta \in \mathcal{L}^-(\Psi)$. However, and unlike the case for random worlds, for random structures once we have equality in the language, a single unary predicate suffices in order to have this probability assume infinitely many values.

THEOREM 4.24. *Fix a finite unary vocabulary Ψ that contains at least one unary predicate and rational numbers $0 \le r_1 \le r_2 \le 1$ such that $[r_1, r_2] \ne [0,1]$. For $\varphi, \theta \in \mathcal{L}(\Psi)$, the problem of deciding whether $\Pr_\infty^{s,\Psi}(\varphi \mid \theta) \in [r_1, r_2]$ is PSPACE-hard, even given an oracle that tells us whether the limit is well defined.*

Proof. We first prove the result under the assumption that $\Psi = \{P\}$.

For the case of $r_1 = 0$ and $r_2 < 1$, the result follows trivially from Theorem 4.23. Let $r_1 = q/p > 0$. As for random worlds, we construct formulas $\varphi_{r_1}, \theta_{r_1}$ such that $\Pr_\infty^{s,\{P\}}(\varphi_{r_1} \mid \theta_{r_1}) = r_1$. The formula θ_{r_1} is $\exists x\, P(x) \wedge \exists^{\le p} x\, P(x)$. The formula φ_{r_1} is $\exists x\, P(x) \wedge \exists^{\le q} x\, P(x)$. Clearly, there are p atomic descriptions consistent with θ_{r_1}, among which q are also consistent with φ_{r_1}. Thus, $\Pr_\infty^{s,\{P\}}(\varphi_{r_1} \mid \theta_{r_1}) = q/p = r_1$.

Now, as in Theorem 4.8, let β be a QBF, and define ξ_β as in that proof. As there, $\Pr_\infty^{s,\{P\}}(\xi_\beta \wedge \varphi_{r_1} \mid \theta_{r_1} \wedge \exists x\, \neg P(x))$ is 0 if β is false and r_1 if it is true. Thus, by computing this probability, we can decide the truth of β, proving PSPACE-hardness in this case.

The result in the case that $\Psi \ne \{P\}$ is not immediate as it is for random worlds, since the asymptotic probability in the case of random structures may depend on the vocabulary. We define a formula θ' to be the following conjunction: for each predicate P' in $\Psi - \{P\}$, θ' contains the conjunct $\forall x P'(x)$. If Ψ contains constant symbols, let c be a fixed constant symbol in Ψ. Then θ' also contains the conjunct $\bigwedge_{P' \in \Psi} P'(c)$, and conjuncts $c = c'$ for each constant c' in Ψ. We leave it to the reader to check that for any formulas $\varphi, \theta \in \mathcal{L}(\{P\})$, $\Pr_\infty^{s,\{P\}}(\varphi \mid \theta) = \Pr_\infty^{s,\Psi}(\varphi \mid \theta \wedge \theta')$. ☐

For the case of a finite vocabulary and a bound on the quantifier depth, precisely the same argument as that given for Theorem 4.9 allows us to show the following.

THEOREM 4.25. *Fix $d \ge 0$. For $\varphi, \theta \in \mathcal{L}(\Psi)$ such that $d(\varphi), d(\theta) \le d$, we can effectively construct a linear time algorithm that decides if $\Pr_\infty^{s,\Psi}(\varphi \mid \theta)$ is well defined and computes it if it is.*

We now drop the assumption that we have a fixed finite vocabulary. As we previously discussed, there are at least two distinct interpretations for asymptotic conditional probabilities in this case. One interpretation of "infinite vocabulary" views Ω as a potential or background vocabulary, so that every problem instance includes as part of its input the actual finite subvocabulary that is of interest. So although this subvocabulary is finite, there is no bound on its possible size. The alternative is to interpret infinite vocabularies more literally, using the limit process explained in §2.3. As we mentioned, for random worlds the two interpretations are equivalent. However, this is not the case for random structures, where the two interpretations may give different answers. In fact, from Corollary 2.9, it follows that the random-structures method reduces to the random-worlds method under the second interpretation. Thus, the complexity results are the same for random worlds and random structures under this interpretation. As we already observed, even under the first interpretation, the random-structures method reduces to the random-worlds method if there is a binary predicate in the language. It therefore remains to prove the complexity results for random structures only for the first interpretation, where the vocabulary is considered to be part of the input, under the assumption that the language is unary. As Example 2.4 shows, in this case, the random-worlds method may give answers different from those given by the random-structures method. Nevertheless, as we now show, the same complexity bounds hold for both random worlds and random structures.

As for the case of the finite vocabulary, the lower bounds for computing the probability (Theorem 4.15) and for approximating it (Theorem 4.16) only use formulas in $\mathcal{L}^-(\mathcal{P})$ for some $\mathcal{P} \subset \mathcal{Q}$. Therefore, by Corollaries 3.41 and 3.42, the lower bounds hold unchanged for random structures.

THEOREM 4.26. *For $\Psi \subset \Upsilon$ and $\varphi, \theta \in \mathcal{L}^-(\mathcal{P})$ of depth at least 2, the problem of computing $\Pr^{s,\Psi}_\infty(\varphi \mid \theta)$ is #EXP-hard, even given an oracle for deciding whether the limit exists.*

THEOREM 4.27. *Fix rational numbers $0 \le r_1 \le r_2 \le 1$ such that $[r_1, r_2] \ne [0, 1]$. For $\Psi \subseteq \Omega$ and $\varphi, \theta \in \mathcal{L}^-(\mathcal{P})$ of depth 2, the problem of deciding whether $\Pr^{s,\Psi}_\infty(\varphi \mid \theta) \in [r_1, r_2]$ is both NEXPTIME-hard and co-NEXPTIME-hard, even given an oracle for deciding whether the limit exists.*

It remains only to prove the #EXP upper bound for computing the asymptotic probability.

THEOREM 4.28. *For $\Psi \subseteq \Omega$ and $\varphi, \theta \in \mathcal{L}(\Psi)$, the problem of computing $\Pr^{s,\Psi}_\infty(\varphi \mid \theta)$ is #EXP-easy.*

Proof. We again follow the outline of the proof for the case of random worlds. Recall that in the proof of Theorem 4.13 we construct a Turing machine such that the number of accepting paths of **M** encodes, for each degree δ, $count^\delta(\varphi)$ and $count^\delta(\theta)$. From this encoding we could deduce the maximum degree, and calculate the asymptotic conditional probability. This was accomplished by guessing a model description $\psi \wedge \mathcal{V}$, and branching sufficiently often, according to $\Delta(\psi)$, so that the different counts in the output are guaranteed to be separated. The construction for random structures is identical, except that we guess atomic descriptions ψ rather than model descriptions, and branch according to $\alpha(\psi)$ rather than according to $\Delta(\psi)$. Again, since there are fewer atomic descriptions than model descriptions, and the representation of atomic descriptions is shorter, the resulting Turing machine is less complex, and therefore also in #EXP. ☐

5. Conclusions. In this paper and [23], we have carried out a rather exhaustive study of complexity issues for two principled methods for computing degrees of belief: the random-worlds method and the random-structures method. These are clearly not the only methods that one can imagine for this purpose. However, as discussed in [2] and [22], both methods are often successful at giving answers that are intuitively plausible and which agree with well-known desiderata. We believe this success justifies a careful examination of complexity issues.

Here we have focused on the case where the formula we are conditioning on is a unary first-order formula. As we mentioned in the introduction, in many applications we want to move beyond first order and also allow for statistical knowledge. Both methods continue to make sense in this case. Furthermore, as shown in [33], [3], and [27], for a unary language we can often calculate asymptotic probabilities in the random-worlds method, using a combination of the techniques in this paper and the principle of maximum entropy. Since a lot is already known about computing maximum entropy (for example, [7], [9], [19]), this combination may lead to efficient algorithms for some practical problems.

Acknowledgments. We would like to thank Fahiem Bacchus, with whom we started working on this general area of research. We would also like to thank Ron Fagin and Moshe Vardi for useful comments on a previous draft of this paper. Ron's comments encouraged us to discuss both the random-structures and the random-worlds methods in this paper. We would also like to thank Moshe for pointing out the reference [31]. Finally, we would like to thank Andrew Granville for help with some of the details of Theorem 4.8.

REFERENCES

[1] W. Ackermann, *Solvable cases of the decision problem*, North-Holland, Amsterdam, 1954.
[2] F. Bacchus, A. J. Grove, J. Y. Halpern, and D. Koller, *From statistics to belief*, in Proc. Tenth National Conference on Artificial Intelligence (AAAI '92), 1992, pp. 602–608.
[3] ———, *From statistical knowledge bases to degrees of belief*, Artificial Intelligence, 87 (1996), pp. 75–143.
[4] A. Blass, Y. Gurevich, and D. Kozen, *A zerio-one law for logic with a fixed point operator*, Information and Control, 67 (1985), pp. 70–90.
[5] A. K. Chandra, D. Kozen, and L. J. Stockmeyer, *Alternation*, Journal of the ACM, 28 (1981), pp. 114–133.
[6] A. K. Chandra and P. M. Merlin, *Optimal implementation of conjunctive queries in relational databases*, in Proc. 9th ACM Symp. on Theory of Computing, 1977, pp. 77–90.
[7] P. C. Cheeseman, *A method of computing generalized Bayesian probability values for expert systems*, in Proc. Eighth International Joint Conference on Artificial Intelligence (IJCAI '83), 1983, pp. 198–202.
[8] K. Compton, *0-1 laws in logic and combinatorics*, in Proc. 1987 NATO Adv. Study Inst. on algorithms and order, I. Rival, ed., Reidel, Dordrecht, Netherlands, 1988, pp. 353–383.
[9] W. E. Deming and F. F. Stephan, *On a least squares adjustment of a sampled frequency table when the expected marginal totals are known*, Annals Mathematical Statistics, 11 (1940), pp. 427–444.
[10] K. G. Denbigh and J. S. Denbigh, *Entropy in Relation to Incomplete Knowledge*, Cambridge University Press, Cambridge, U.K., 1985.
[11] B. Dreben and W. D. Goldfarb, *The Decision Problem: Solvable Classes of Quantificational Formulas*, Addison-Wesley, Reading, Mass., 1979.
[12] H. B. Enderton, *A Mathematical Introduction to Logic*, Academic Press, New York, 1972.
[13] R. Fagin, *Probabilities on finite models*, Journal of Symbolic Logic, 41 (1976), pp. 50–58.
[14] ———, *The number of finite relational structures*, Discrete Mathematics, 19 (1977), pp. 17–21.
[15] H. Gaifman, *Probability models and the completeness theorem*, in International Congress of Logic Methodology and Philosophy of Science, 1960, pp. 77–78. This is the abstract of which [16] is the full paper.
[16] ———, *Concerning measures in first order calculi*, Israel Journal of Mathematics, 2 (1964), pp. 1–18.
[17] M. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-completeness*, W. Freeman and Co., San Francisco, 1979.
[18] 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 (1969), pp. 17–28.
[19] S. A. Goldman, *Efficient methods for calculating maximum entropy distributions*, master's thesis, MIT EECS Department, 1987.
[20] E. Grandjean, *Complexity of the first-order theory of almost all structures*, Information and Control, 52 (1983), pp. 180–204.
[21] 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, pp. 294–305.
[22] ———, *Random worlds and maximum entropy*, Journal of A.I. Research, 2 (1994), pp. 33–88.
[23] ———, *Asymptotic conditional probabilities: the non-unary case*, Journal of Symbolic Logic, 61 (1996), pp. 250–275.
[24] I. Hacking, *The Emergence of Probability*, Cambridge University Press, Cambridge, U.K., 1975.
[25] E. T. Jaynes, *Where do we stand on maximum entropy?*, in The Maximum Entropy Formalism, R. D. Levine and M. Tribus, eds., MIT Press, Cambridge, Mass., 1978, pp. 15–118.
[26] J. M. Keynes, *A Treatise on Probability*, Macmillan, London, 1921.
[27] D. Koller, *From Knowledge to Belief*, PhD thesis, Stanford University, Dept. of Computer Science, 1994.
[28] J. v. Kries, *Die Principien der Wahrscheinlichkeitsrechnung und Rational Expectation*, Mohr, Freiburg, 1886.
[29] H. R. Lewis, *Unsolvable Classes of Quantificational Formulas*, Addison-Wesley, New York, 1979.
[30] ———, *Compelexity results for classes of quantificational formulas*, Journal of Computer and System Sciences, (1980), pp. 317–353.
[31] M. I. Liogon'kiĭ, *On the conditional satisfiability ratio of logical formulas*, Mathematical Notes of the Academy of the USSR, 6 (1969), pp. 856–861.

[32] J. LYNCH, *Almost sure theories*, Annals of Mathematical Logic, 18 (1980), pp. 91–135.
[33] J. B. PARIS AND A. VENCOVSKA, *On the applicability of maximum entropy to inexact reasoning*, International Journal of Approximate Reasoning, 3 (1989), pp. 1–34.
[34] S. J. PROVAN AND M. O. BALL, *The complexity of counting cuts and of computing the probability that a graph is connected*, SIAM Journal on Computing, 12 (1983), pp. 777–788.
[35] D. ROTH, *On the hardness of approximate reasoning*, in Proc. Thirteenth International Joint Conference on Artificial Intelligence (IJCAI '93), 1993, pp. 613–618.
[36] G. SHAFER, *Personal communication*, 1993.
[37] L. J. STOCKMEYER, *The polynomial-time hierarchy*, Theoretical Computer Science, 3 (1977), pp. 1–22.
[38] L. G. VALIANT, *The complexity of computing the permanent*, Theoretical Computer Science, 8 (1979), pp. 189–201.
[39] ———, *The complexity of enumeration and reliability problems*, SIAM Journal on Computing, 8 (1979), pp. 410–421.
[40] R. L. VAUGHT, *Applications of the Lowenheim-Skolem-Tarski theorem to problems of completeness and decidability*, Indagationes Mathematicae, 16 (1954), pp. 467–472.