Random theories

Nikolay Bazhenov[1], Samuel Braunfeld[2], and Kyle Gannon[3]

[1] Sobolev Institute of Mathematics, Novosibirsk, Russia
nickbazh@yandex.ru
[2] Computer Science Institute of the Czech Academy of Sciences, Prague, Czechia
braunfeld@cs.cas.cz
[3] Beijing International Center for Mathematical Research, Beijing, China
kgannon@bicmr.pku.edu.cn

Abstract

We consider random theories, i.e. Borel probability measures on the space of complete theories in a fixed language. We particularly focus on the theories obtained by taking a Bernoulli unary expansion of a countable structure.

1 Introduction

There have been a variety of results connecting model theory and random labeled structures, i.e. Borel probability measures on countable labeled L-structures for some fixed language L, often focusing on exchangeable random structures such as graphons. A key point of model theory it is often helpful to study theories rather than structures, and so we propose studying random theories, i.e. Borel probability measures on the space of complete (first-order) L-theories. One benefit of this move is that there are settings that lead to a well-defined random theory, even if they do not lead to any well-defined random structure.

One natural way to arrive at a random theory is to take a probability distribution on a class of finite L-structures that has a convergence law, which generalizes 0-1 laws by just requiring that the probability of each sentence converges, rather than that it converges to 0 or 1. This is a quite classical topic with an enormous literature. Such random theories will be pseudofinite, in a sense that we define later. Another natural way to arrive at a random theory is to start with a (random or deterministic) countable labeled L-structure, perform a random process, and take the random theory of the resulting random structure. One simple case of this that we will focus on is to start with a structure M, expand by a new unary relation P, and then independently add each point to P with probability $1/2$; we will refer to the resulting random theory as the *Bernoulli theory* of M.

2 Preliminaries

We recall that the space T_L of complete L-theories is a totally disconnected compact Hausdorff space, with basic clopen sets of the form $[\varphi] \coloneqq \{ T \mid \varphi \in T \}$ where φ is a sentence.

Definition 1. A *random theory* is a regular Borel probability measure on T_L.

Definition 2. Let M be a countable labeled L-structure, and where L does not contain P. The *Bernoulli theory of* M is the random $(L \cup \{ P \})$-theory obtained by first taking the random structure ν given by expanding M by a Bernoulli coloring with $p = 1/2$ and interpreting this as the new unary relation P. Then we consider the map sending a labeled structure to its theory, and take the pushforward of ν under this map.

We remark that although the random structure ν above is dependent on the enumeration of M, the Bernoulli theory is not.

Also, one could consider the *Bernoulli substructure* of a relational structure, by taking substructure induced by the new unary relation P. All the questions we consider have parallels for the Bernoulli substructure, although not necessarily parallel answers.

Although we will not make use of it, we mention the following theorem, showing that in some sense all the randomness in the Bernoulli theory of M is contained in $\mathrm{acl}(\varnothing)$. The first part is essentially proved by showing that the action of $Aut(M)$ on the space of unary expansions is ergodic.

Theorem 2.1 (Mycielski; see [5]). *Let M be a countable structure such that $Aut(M)$ has no finite orbits. Then in the Bernoulli theory of M, every L_P-sentence has measure 0 or 1. Furthermore, the set of expansions that satisfy a given L_P-sentence is either meager or comeager.*

In fact, the theorem holds also for uniformly random expansions in higher arities. In [5], Lynch identifies a general condition for when these two complete L_P-theories agree.

3 The Bernoulli theory is not preserved by elementary equivalence

It is often convenient to be able to work with a "monster model" of a theory with good saturation and homogeneity properties, rather than having to consider each model individually. In this section, we see that we cannot in general do this with the Bernoulli theory. We give an example due to Gregory Cherlin (personal communication) of (NIP) structures $M \equiv N$ where taking the Bernoulli theory gives rise to distinct theories; in particular, there will be a sentence φ such that the Bernoulli theory of M does not satisfy φ, while the Bernoulli theory of N almost surely satisfies φ.

Example 1. We first define the structure M. It has domain $\{(i,j) \mid j < i \in \mathbb{N}\}$. The language is (E, s_1, s_2, f) where E, s_1, s_2 are a binary relations, and f is (the graph of) a partial binary function. Roughly, the structure M will have an equivalence relation with a class of each finite size, each class will be equipped with a successor relation and there will be the natural successor relation on classes, and there is a "translation" function describing how to consecutively embed one class in another. We interpret E as the equivalence relation of agreement in the first coordinate, $s_1((i,j),(i',j'))$ holds when $i' = i+1$, $s_2((i,j),(i',j'))$ holds when $j' = j+1$ and $i' = i$, and f given by $f((i,j),(i',j')) = (i, j+j')$ (which is only defined so long as $j+j' < i$). Note that M is definable in Pressburger arithmetic, and so NIP. Let N be a countable elementary extension of M. Note that N will have an infinite E-class, and the theory ensures that for every s_2-initial point c, we have $f(x,c) = x$ and that f respects s_2, i.e. for every d, if $f(x,d) = y$ and $s_2(d,d')$ and $s_2(y,y')$ then $f(x,d') = y'$. Thus f also describes consecutive embeddings of the finite E-classes into the infinite ones, and of s_2-closed initial segments of one infinite E-class into another E-class. Also note that the set of finite E-classes is closed under the successor relation s_1.

Now consider the expanded language L_P with a new unary relation P. First, we write a formula $\varphi(x,y)$ that says that the equivalence class x/E consecutively embeds into y/E via f in a way that respects P:

$$\varphi(x,y) \coloneqq \exists y'(E(y,y') \wedge \forall x'(E(x,x') \to \exists z(z = f(y',x') \wedge (P(z) \leftrightarrow P(x')))$$

Now let ψ be the sentence that says there exists a c such that $\varphi(x,c)$ is closed under s_1. No c such that c/E is finite can be a witness for ψ, and so the Bernoulli theory of M gives ψ measure 0. On the other hand, in the Bernoulli theory of N, if c/E is infinite then $\varphi(x,c)$ almost surely defines the union of the finite classes, so c is almost surely a witness for ψ, and so ψ receives measure 1.

4 Bernoulli theories and dividing lines

There are many results on when various sorts of unary expansions preserve model-theoretic dividing lines. For a property P, we will say a theory is *monadically* P if every unary expansion of the theory has P. There are characterizations of when theories are monadically NFCP (short for "not the finite cover property"), monadically stable, or monadically NIP (short for "not the independence property"). On the other hand, if a theory has the independence property, then a monadic expansion can define a model of ZFC, and so it cannot be monadically P for any reasonable P. We will say a structure M is *Bernoulli-monadically* P if the Bernoulli theory of M almost surely has P. We will give examples showing that for NFCP/stable/NIP, Bernoulli-monadically P is strictly between monadically P and P. For simplicity, it seems possible that every simple structure M is Bernoulli-monadically simple.

Example 2. NFCP We recall that a theory T is NFCP if it is stable and T^{eq} eliminates $\exists^\infty$. In particular, every ω-categorical stable theory is NFCP.

For a structure that is Bernoulli-monadically NFCP but not monadically NFCP, one can take the an equivalence relation with infinitely many classes, all infinite. The theory is not monadically NFCP because there is a unary expansion that cuts out growing finite subclasses. But a Bernoulli coloring almost surely picks an infinite-coinfinite subset from each class, which remains stable and ω-categorical.

For a structure that is NFCP but not even Bernoulli NIP, consider a 1-subdivided complete graph. This is definable in the theory of equality. After a Bernoulli coloring, the structure will almost surely define the random graph via the formula $\varphi(x,y) := \exists z(P(z) \wedge E(x,z) \wedge E(y,z))$.

Stability/NIP For a structure that is stable but not Bernoulli-monadically NIP, the 1-subdivided complete graph again works.

For a structure that is not monadically stable but is Bernoulli-monadically stable, consider a 1-subdivded complete graph, but with infinitely many copies of all the subdivision paths. More formally, the domain is $\mathbb{N} \cup \mathbb{N}^3$ where there is an edge between i and (j,k,l) if $i = j$ or $i = k$. Then almost surely a Bernoulli coloring picks out an infinite-coinfinite subset of $\mathbb{N}$, and for each $i, j \in \mathbb{N}$ picks out an infinite-coinfinite subset of $\{\, (i,j,k) \mid k \in \mathbb{N}\,\}$.

For simplicity, we have the following question.

Question 1. Is every simple structure Bernoulli-monadically simple*?*

We remark that [3, Corollary 2.8] provides some indication that a positive answer is not so unreasonable.

5 Pseudofiniteness

A benefit of random theories over random labeled structures is that random theories of finite and infinite structures naturally live in the same space. We say a random theory is *uniformly finite* if every theory in its support has a finite model.

Definition 3. A random theory μ is *pseudofinite* if whenever $\mu(\varphi) > 0$, φ has a finite model.

It then follows that pseudofinite random theories are the topological closure of uniformly finite random theories, and that a random theory is pseudofinite if and only if it is an ultralimit of uniformly finite random theories (each with finite support).

Question 2. Let M be pseudofinite. Is the Bernoulli theory of M pseudofinite? In particular, can this be witnessed by taking Bernoulli colorings of the witnesses for the pseudofiniteness of M?

As mentioned before, the limiting random theories of classes of finite structures with a first-order convergence law give many examples of pseudofinite random theories. For a random theory derived from a natural infinitary random process that is not pseudofinite, we mention the *countable threshold graph.*

Definition 4. *The* countable threshold graph *is a measure on the space of countable labeled graphs defined as follows: for each natural number n, independently and with probability 1/2 we either connect it to all preceding natural numbers or to no preceding natural numbers.*

One can definably recover the partition (X, Y) of the vertices into those that were and were not connected to the preceding natural numbers (see the formulas for W_1 and W_2 in [1, Proposition 4.2]). Then one can define a partial order on Y by $y_1 \leq y_2 \iff \exists(z \in X)(R(z, y_1) \wedge \neg R(z, y_2)$. This partial order almost surely does not have a maximal element, and so the theory is not pseudofinite. One can show, for appropriate definitions of stability and NIP for random theories, that the theory is NIP and unstable.

6 Models of random theories

The following seems like a natural definition for a model of a random theory. In order to handle theories like a weighted sum of two theories where one has only finite models and one has only infinite models, we must allow for our models to range over $\mathcal{L}$-structures of different sizes.

Definition 5. *Fix a language $\mathcal{L}$ and a random $\mathcal{L}$-theory μ. Let π be the function mapping an $\mathcal{L}$-structure to its theory. A* model of μ *is a Borel probability measure ν on a space $X = \bigsqcup_{i\in\omega+1, i>0} X_i$ where X_i is the space of $\mathcal{L}$-structures with domain i (equipped with the usual topology generated by basic clopen sets consisting of the structures satisfying a quantifier-free formula on a particular tuple) such that the pushforward $\pi_*\nu = \mu$. (Note that the restriction of π to X is Borel, by [4, Proposition 16.7].)*

The restriction to spaces of countable structures may not be necessary, but it is unclear what the topology should be on spaces of uncountable structures. If we keep the topology as above, then the set of structures satisfying a given sentence need not be Borel. If we drop the quantifier-free condition from the formulas generating the topology, then the space is no longer compact.

Using a result about pulling back measures between standard Borel spaces, we can show such models exist.

Theorem 6.1 (Completeness for random theories)**.** *Fix a countable language $\mathcal{L}$. Let μ be a random $\mathcal{L}$-theory. Then μ has a model.*

Proof. Let X be as in Definition 5, so X is a standard Borel space. By the ordinary completeness and downwards Löwenheim-Skolem theorems, the map π sending an $\mathcal{L}$-structure to its theory induces a (Borel) surjection from X onto the space of $\mathcal{L}$-theories. By [2, Theorem 9.1.5], there is a measure ν on X such that $\pi_* \nu = \mu$. □

Question 3. *How much further can the model theory of random theories be developed? Are there reasonable notions of, for example, definable sets, types, and elementary extensions? Are there reasonable notions of uncountable models, and do the Löwenheim-Skolem theorems hold?*

References

[1] Nathanael Ackerman, Cameron Freer, Kyle Gannon, James E Hanson, and Rehana Patel. Generic sampling and invariant measures on the space of k-uniform hypergraphs. *arXiv preprint arXiv:2510.17090*, 2025.

[2] V. I. Bogachev. *Measure theory*, volume 2. Springer, 2007.

[3] Zoé Chatzidakis and Anand Pillay. Generic structures and simple theories. *Annals of Pure and Applied Logic*, 95(1-3):71–92, 1998.

[4] Alexander Kechris. *Classical descriptive set theory*. Springer Science & Business Media, 2012.

[5] James F Lynch. Almost sure theories. *Annals of Mathematical Logic*, 18(2):91–135, 1980.